

Министерство образования и науки Российской Федерации

Калужский государственный университет  
им. К.Э. Циолковского

**Е.В. Леонова**

**КАЧЕСТВЕННЫЕ И  
КОЛИЧЕСТВЕННЫЕ МЕТОДЫ  
ИССЛЕДОВАНИЯ В ПСИХОЛОГИИ**

*Учебник*

**Ай Пи Эр Медиа  
Саратов • 2017**

УДК 159.9 : 378

ББК 88

Л 47

**Леонова, Е. В.**

**Л47** Качественные и количественные методы исследования в психологии [Электронный ресурс]: учебник / Е. В. Леонова; Калужский государственный университет им. К.Э. Циолковского. — Электрон. дан. и прогр. (8 Мб). — Саратов: Ай Пи Эр Медиа, 2017. — 411 с.

Учебник разработан в соответствии с Федеральным Государственным образовательным стандартом по дисциплинам «Психодиагностика», «Математические методы в психологии», «Многомерный статистический анализ данных», «Качественные и количественные методы исследования в психологии». В учебнике изложены основные темы указанных дисциплин, рассмотрены основные методы психологической оценки различных уровней формализации, представлена их классификация. Раскрыты основные принципы конструирования инструментов исследования в психологии. Подробно и доступным языком изложены методы математико-статистического анализа данных психологического исследования.

Учебник предназначен для студентов (бакалавров, специалистов и магистров) психологических направлений подготовки, аспирантов и соискателей. Учебник может оказаться полезным для преподавателей высших учебных заведений и научных работников.

Рецензент: д.психол.н., проф. В.Ф. Енгальчев

*Издается по решению Ученого Совета КГУ им. К.Э.Циолковского*

***Учебное электронное издание***

*Для создания электронного издания использовано:*

Приложение pdf2swf из ПО Swftools, ПО IPRbooks Reader,  
разработано на основе Adobe Air.

Подписано к использованию 31.10.2017.

ISBN 978-5-4486-0260-3

© Леонова Е. В., 2017

© КГУ им. К.Э.Циолковского, 2017

© ООО «Ай Пи Эр Медиа», 2017



## Оглавление

<b>Предисловие.....</b>	<b>5</b>
<b>Глава 1. Измерение и оценка в психологии .....</b>	<b>7</b>
1.1. Понятия методологии, метода и методики исследования в психологии.....	7
1.2. Психодиагностический метод измерения и оценки.....	18
1.3. Классификация методов психологической оценки .....	27
1.4. Шкалы измерения и экспериментальная выборка .....	30
<b>Глава 2. История развития методов оценивания в психологии и образовании .....</b>	<b>35</b>
2.1. История развития методов психологической диагностики и оценки.....	35
2.2 История развития системы оценивания в образовании .....	56
<b>Глава 3. Строгоформализованный метод психологической оценки .....</b>	<b>62</b>
3.1. Классификация строгоформализованных методик.....	62
3.2. Метрические характеристики строгоформализованных методик .....	86
3.3. Этапы конструирования строгоформализованных методик.....	110
3.4 Адаптация зарубежных строгоформализованных методик..	126
<b>Глава 4. Малоформализованные методы психологической оценки .....</b>	<b>130</b>
4.1. Метод наблюдения.....	130
4.2. Метод анкетного опроса .....	133
4.3. Беседа, интервью .....	148
4.4. Метод экспертных оценок.....	154
4.4. Контент-анализ .....	159
<b>Глава 5. Психологическая оценка компетенций в образовательном процессе .....</b>	<b>174</b>
5.1. Компетентностный подход к результатам высшего профессионального образования .....	174

5.2. Критерии и показатели сформированности общекультурных и профессиональных компетенций студентов .....	185
<b>Глава 6. Методы первичной обработки данных психологического исследования .....</b>	<b>190</b>
6.1. Первичная обработка данных психологического исследования .....	190
6.2. Графическое представление данных .....	199
6.3. Числовые характеристики множества данных .....	215
6.4. Распределение признака.....	227
<b>Глава 7. Методы статистического анализа данных .....</b>	<b>237</b>
7.1. Основные понятия статистической обработки данных .....	239
7.2. Статистические критерии различий и изменений.....	251
7.3. Статистические критерии согласия распределений .....	280
7.4. Меры связи .....	289
7.5. Многофункциональные статистические критерии .....	301
<b>Глава 8. Обзор методов многомерного анализа данных.....</b>	<b>309</b>
8.1. Регрессионный анализ.....	310
8.2. Дисперсионный анализ .....	316
8.3. Группировка испытуемых методами дискриминантного и кластерного анализа.....	323
8.4. Группировка переменных с помощью методов факторного анализа .....	330
<b>Глава 9. Использование методов измерения и оценки в психологических исследованиях .....</b>	<b>337</b>
9.1. Правила составления пакета методик психологического исследования.....	337
9.2 Пример комплексного использования методов измерения и оценки.....	344
<b>Литература .....</b>	<b>377</b>
<b>Приложения .....</b>	<b>386</b>

## Предисловие

Мы живём в интенсивно меняющемся мире, когда за короткие промежутки времени происходят кардинальные изменения в научно-технологическом и в социально-экономическом развитии. Современное общество отличается неустойчивостью социальной, экономической, идеологической обстановки, сменой нравственных ориентиров, и учёт психологических ресурсов людей повышает результативность их деятельности во всех областях социальной практики, таких как образование, профессиональная деятельность, семейная жизнь. В этих условиях возросла потребность общества в квалифицированных психологах, владеющих современными психодиагностическими технологиями, способных разрабатывать инструменты измерения и оценки индивидуально-психологических особенностей личности в различных сферах деятельности.

Наиболее высокая потребность в разработке методов и технологий оценки наблюдается сегодня в сфере образования. В соответствии с новыми образовательными стандартами результатом образования становятся знания, умения, навыки и освоенные компетенции, оценить которые возможно лишь при использовании достижений современной психологической науки, в том числе, современной психологической диагностики. Ключевым фактором повышения эффективности образования на всех уровнях является разработка и внедрение в образовательный процесс психологических технологий развития и оценки сформированности компетенций, определяющими факторами эффективности которых являются научная обоснованность и системность.

В настоящем учебнике методы психологической оценки рассматриваются в контексте общей идеологии научного исследования. Учебник разработан в соответствии с Федеральным Государственным образовательным стандартом по дисциплинам «Психодиагностика», «Математические методы в психологии», «Много-

мерный статистический анализ данных», «Качественные и количественные методы психологического исследования». Эмпирический материал, на основе которого разработаны примеры заданий, получен в результате исследований, осуществлённых под руководством автора учебника.

В учебнике рассмотрены основные методы психологической оценки различных уровней формализации, представлена их классификация. Раскрыты основные принципы конструирования инструментов психологической оценки, особое внимание уделяется методам оценки сформированности личностных компетенций. Подробно и доступным языком изложены методы математико-статистического анализа данных психологического обследования.

Содержание учебника условно можно разделить на две части. Первая часть (главы 1-5) посвящена методологии, истории, классификации, основным принципам построения измерительных инструментов в психологии. Во второй части (главы 6-8) раскрываются методы математико-статистического анализа результатов измерения в психологии. Заключительная глава посвящена комплексному использованию методов измерения и оценки в психологических исследованиях

## **Глава 1.**

### **Измерение и оценка в психологии**

#### **1.1. Понятия методологии, метода и методики исследования в психологии**

Метод – это способ научного познания объекта или практической деятельности, реализующий познавательную позицию субъекта к объекту исследования [21, с. 306]. Выбор метода исследования отражает систему концептуальных принципов исследователя, представляющих собой методологическую основу исследования. Другими словами, в методах конкретизируется методология научного исследования. Для определения роли и места методов измерения и оценки в методологии исследований в области психологии рассмотрим общие подходы к научному исследованию.

Согласно логическому словарю-справочнику Н.И.Кондакова исследование – это процесс научного изучения какого-либо объекта (предмета, явления – материального или идеального) с целью выявления его закономерностей возникновения, развития и изменения и преобразования его в интересах общества. «Всякое подлинное исследование – отмечает Н.И.Кондаков, – есть единство накопленного предшествующего опыта, имеющихся знаний, применения соответствующих инструментов и орудий и методов, способов подхода к изучаемому объекту» [33, с.217]. Итогом исследования должно быть получение новых научных знаний – объективной истины, т.е. соответствия вновь сформулированного знания с действительным состоянием объекта, а также намеченных программой исследования практических результатов. Непременным компонентом нового знания должно быть понимания закономерностей дальнейшего развития исследуемого объекта.

Научное знание определяется следующими характеристиками (Б.Веймер, [120]):

- научное знание основано на твёрдых эмпирических данных;
- теории выводятся из фактов (и, следовательно, вторичны по отношению к ним);
- наука развивается посредством постепенного накопления фактов;
- поскольку факты лежат в основе нашего знания, они независимы от теорий и имеют самостоятельное значение;
- теории (или гипотезы) логически выводятся из фактов посредством рациональной индукции;
- теории (или гипотезы) принимаются или отвергаются исключительно в зависимости от их способности выдержать проверку экспериментом [цит. по 108, с. 21-22].

Указанные характеристики научного познания считаются нормативными и закреплены в стандартной структуре диссертаций, научных статей, дипломных и курсовых работ. К.Поппером [67, 68] были определены следующие составляющие процесса научного исследования:

- А. Выдвижение гипотезы (гипотез).
- Б. Планирование исследования.
- В. Проведение исследования.
- Г. Интерпретация данных.
- Д. Опровержение или подтверждение гипотезы (гипотез).
- Е. В случае опровержения старой – формулирование новой гипотезы (гипотез).

В соответствии с классификацией В.Н.Дружинина [21] существуют следующие **типы научных исследований**:

1. Поисковые исследования. Это попытка решения проблемы, которую никто не ставил или не решал подобным методом. Научные работы такого рода направлены на получение принципиально новых результатов в малоисследованной области.
2. Критические исследования. Они проводятся в целях опровержения существующей теории, модели, гипотезы, закона и пр.

или для проверки того, какая из двух альтернативных гипотез точнее прогнозирует реальность. Критические исследования проводятся в тех областях, где накоплен богатый теоретический и эмпирический запас знаний и имеются апробированные методики для осуществления эксперимента.

3. Уточняющие исследования. Их цель – установление границ, в пределах которых теория предсказывает факты и эмпирические закономерности. Обычно, по сравнению с первоначальным экспериментальным образцом, изменяются условия проведения исследования, объект, методика. Тем самым регистрируется, на какую область реальности распространяется полученное ранее теоретическое знание.

4. Воспроизводящее исследование. Его цель – точное повторение эксперимента предшественников для определения достоверности, надежности и объективности полученных результатов. Результаты любого исследования должны повториться в ходе аналогичного эксперимента, проведенного другим научным работником, обладающим соответствующей компетенцией. Поэтому после открытия нового эффекта, закономерности, создания новой методики и т. п. возникает лавина воспроизводящих исследований, призванных проверить результаты первооткрывателей.

Научный результат должен быть инвариантным относительно пространства, времени, типа объектов и субъектов исследования, то есть объективным. Вместе с тем научный результат в психологических исследованиях зависит от способа познания, т.е. метода: факт есть функция от свойств объекта и операции с ним.

Согласно А.И.Ракитову основными компонентами (этапами) научного исследования считаются:

- постановка задачи;
- предварительный анализ имеющейся информации, условий и методов решения задач данного класса;
- формулировка исходных гипотез;
- планирование и организация эксперимента;
- проведение эксперимента;



- анализ и обобщение полученных результатов;
- проверка исходных гипотез на основе полученных фактов;
- окончательная формулировка новых фактов и законов, получение объяснений или научных предсказаний [33, с. 218].

Указанные компоненты научного исследования, являясь общеметодологическими, имеют свою специфику для различных наук. Любое явление можно рассматривать с позиций разных точек зрения. При разработке дизайна исследования и анализе его результатов главное – объяснить, в какой парадигме изучаемое явление рассмотрено, а также – как его можно объяснить с позиций других парадигм.

Научное исследование в психологии включает в себя следующие обязательные этапы.

### **I этап. Формулировка научного аппарата исследования.**

*1. Актуальность и проблема исследования.* Для доказательства актуальности достаточно сформулировать проблемную ситуацию. Проблема в науке – это противоречивая ситуация, требующая своего разрешения. Формулировка проблемы означает осознание противоречия, в разрешение которого вносит вклад проведение научного исследования. Пример: острая необходимость в творчески мыслящих специалистах для инновационных сфер производства и ограниченность возможностей общего и профессионального образования в развитии креативности школьников и студентов; следствие – необходимость разработки и внедрения в образовательный процесс методов развития креативности обучающихся, сочетания формального, неформального и информального образования.

*2. Цель исследования.* Цель – это тот конечный результат, который должен быть достигнут в результате проведения исследования, например: разработать методы развития творческого мышления школьников и студентов в условиях формального, неформального, информального образования. Формулировка цели исследования чётко соответствует теме и предмету исследования.

*3. Объект и предмет исследования.* Объект и предмет исследования соотносятся между собой как общее и частное. Объект пси-



хологического исследования – это психический процесс или явление, избранное для исследования (например, творческое мышление школьников и студентов). Предмет исследования – это часть объекта, призма, через которую изучается объект. Применительно к рассматриваемому примеру можно предложить следующие формулировки предмета исследования.

1. Особенности развития творческого мышления школьников и студентов.
2. Показатели креативности школьников и студентов.
3. Неформальное образование школьников как фактор развития творческого мышления.
4. Возможности образовательной среды школы и вуза по развитию творческого мышления обучающихся.
5. Активные методы развития творческого мышления студентов технического вуза.

И так далее. Именно предмет исследования определяет тему работы и зачастую почти повторяет название темы.

4. *Задачи исследования* – последовательные этапы достижения цели исследования, организации и проведения исследования с начала до конца. Как правило, в первой задаче рассматривается объект исследования, а во второй – предмет, либо «та призма, через которую рассматривается объект».

**II этап. Теоретический анализ информации по изучаемой проблеме.**

На этом этапе исследователь изучает результаты проведённых исследований по определенной им проблеме. Теоретико-аналитический этап исследования обязательно включает анализ существующих методов решения данной и аналогичных проблем. По результатам теоретического анализа определяются:

- 1) **теоретическая основа исследования** – труды учёных, внесших вклад в решение проблемы, публикации в научных журналах, инструктивные и справочные материалы, нормативные документы, законодательные акты и т.п.;

- 2) **методологическая основа исследования** – система концептуальных положений и принципов, на основе которых строится модель и формулируется гипотеза исследования.

Исследовательскую модель можно назвать центральным системообразующим элементом всего исследования. В ней аккумулируются выводы по результатам теоретического анализа исследуемой проблемы, определяющие методологическую основу исследования, на основе этих выводов определяются методические подходы к решению проблемы, обосновывается программа экспериментальной работы и её предположительный результат. После того, как выстроена модель, формулируется гипотеза исследования.

Гипотеза – это утверждение предположительного характера, которое в результате исследования должно быть подтверждено или опровергнуто. Гипотеза должна быть сформулирована таким образом, чтобы, с одной стороны, не являться очевидным утверждением, а с другой стороны, иметь потенциальную возможность подтверждения или опровержения в ходе исследования. Гипотеза должна удовлетворять следующим требованиям: а) формулировка гипотезы должна быть максимально точной и недвусмысленной, в ней не должно содержаться неопределённых и неоднозначно трактуемых терминов и понятий; б) формулировка гипотезы и содержание экспериментальной части научного исследования должны находиться в чётком соответствии.

Методологическая основа, модель и гипотеза исследования определяют выбор методов исследования.

### **III этап. Определение методов исследования.**

По степени общности методы исследования делятся на всеобщие, общие и частные.

*Всеобщие методы* – это способы познания, которые могут быть использованы в любой области знаний и любой наукой. Они позволяют определить общие отношения, закономерности и свойства изучаемого объекта или деятельности. К всеобщим относятся методы научного анализа и синтеза, классификации, моделирования [15, 55].

*Общими* называются методы, которые могут применяться в отдельных областях знания и использоваться многими (не всеми) науками. К их числу относятся наблюдение, эксперимент, интерпретационные методы (в том числе, методы представления и математико-статистической обработки данных). [96, 21].

*Частными* (конкретно-научными) называются методы, применяемые в отдельных областях знания, группах смежных наук. К методам психологических исследований относятся методы психологической оценки, психодиагностический метод, интроспекция, социометрический метод и другие методы.

В свою очередь, методы исследования конкретизируются в различных вариантах процедур, операций, правил сбора и анализа данных об изучаемых объекте и предмете.

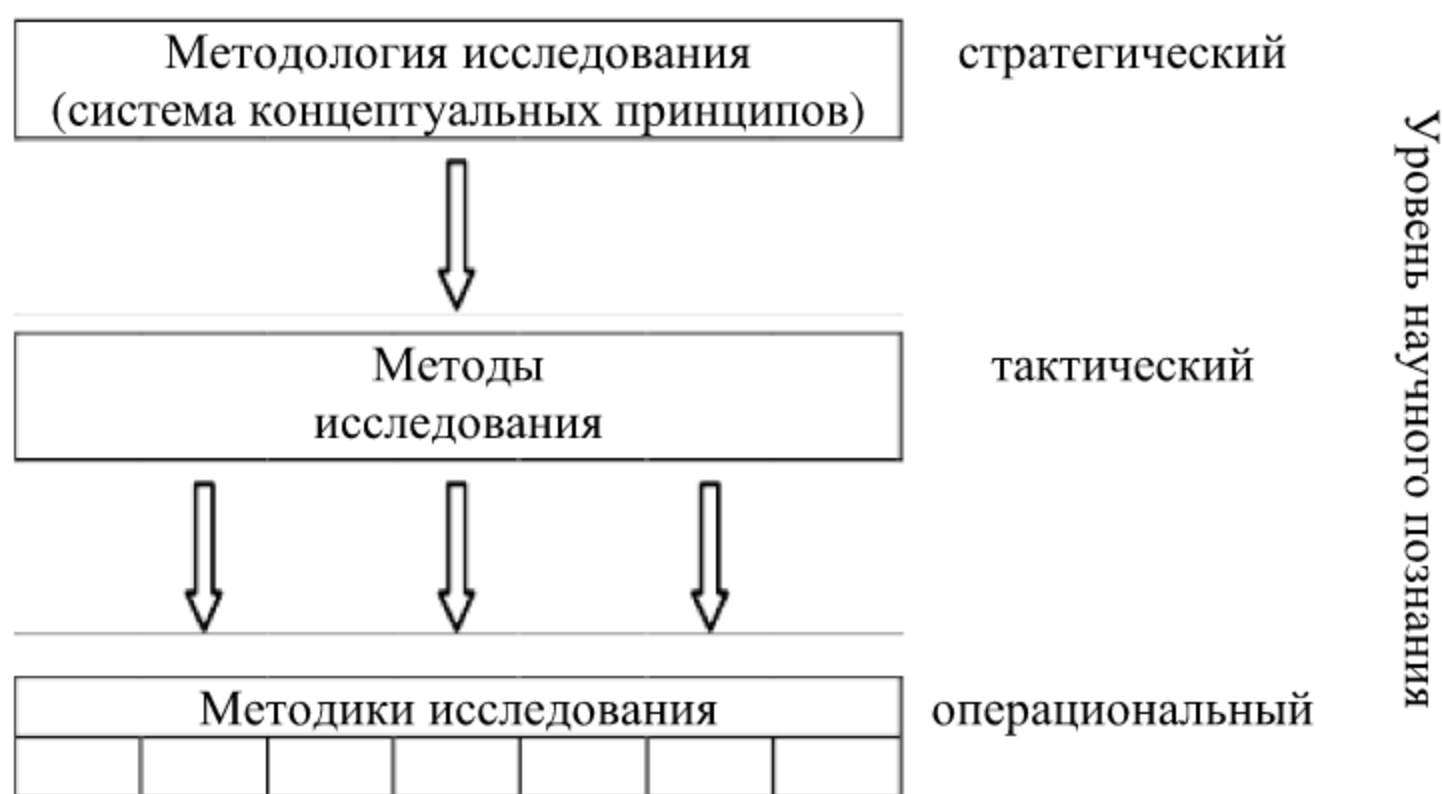
Исследовательские технологии, в которых конкретизируется метод исследования, называются методиками. По определению В. Н. Дружинина методика – это система и последовательность действий исследователя, средств (инструментов, приборов, обстановки), позволяющая решить исследовательскую задачу. С помощью методики фиксируют характеристики поведения и воздействуют на объект.

Итак, категория «метод» тесно связана с некоторым множеством понятий, раскрывающих содержание познавательного процесса. Ближайшими из этих понятий являются «методология» и «методика». Первое понятие отражает стратегический уровень научного познания, второе – оперативный. Метод в этой системе олицетворяет промежуточный, тактический уровень познания (рис. 1.1).

В соответствии с теорией Б.Г.Ананьева, методы психологического исследования являются системами операций с психологическими объектами и вместе с тем являются гносеологическими объектами самой психологической науки.

Методы психологического исследования, по Б.Г. Ананьеву, подразделяются на **организационные**: сравнительный, лонгитюдный, комплексный; **эмпирические**: наблюдение (самонаблюдение), эксперимент (лабораторный, полевой, естест-

венный), психофизиологические, анализ продуктов и процессов деятельности (праксиметрические методы), моделирование, биографический метод; **методы обработки данных**: методы математической статистики, качественного описания; **интерпретационные** методы: генетические (фило- и онто-), структурные (классификация, типологизация).



**Рис. 1.1. Методология, методы и методики исследования**

В.В.Никандровым на основе классификации методов Б.Г.Ананьева разработана следующая классификация методов психологического исследования [55]:

#### А. Организационные методы (подходы)

- а) сравнительный.
- б) лонгитюдный.
- в) комплексный.

#### Б. Эмпирические методы

##### 1. Обсервационные (наблюдение):

- а) объективное наблюдение;
- б) самонаблюдение (интроспекция).

2. Вербально-коммуникативные методы:
  - а) беседа;
  - б) опрос (интервью и анкетирование).
3. Экспериментальные методы:
  - а) лабораторный эксперимент;
  - б) естественный эксперимент;
  - в) формирующий эксперимент.
4. Психодиагностические методы:
  - а) психодиагностические тесты;
  - б) психосемантические методы;
  - в) психомоторные методы;
  - г) методы социально-психологической диагностики личности.
5. Психотерапевтические методы;
6. Методы изучения продуктов деятельности:
  - а) метод реконструкции;
  - б) метод изучения документов (архивный метод);
  - в) графология.
7. Биографические методы;
8. Психофизиологические методы:
  - а) методы изучения работы вегетативной нервной системы;
  - б) методы изучения работы соматической нервной системы;
  - в) методы изучения работы центральной нервной системы.
9. Праксиметрические методы:
  - а) общие методы исследования отдельных движений и действий;
  - б) специальные методы исследования трудовых операций и деятельности.
10. Моделирование;
  - а) моделирование психики;
  - б) психологическое моделирование.
11. Специфические методы отраслевых психологических наук.

## В. Методы обработки данных

- Количественные методы;
- Качественные методы.

#### Г. Интерпретационные методы (подходы)

- Генетический.
- Структурный.
- Функциональный.
- Комплексный.
- Системный.

Анализ представленной классификации методов психологического исследования показывает, что практически все эмпирические методы и методы обработки данных можно отнести к измерению и оценке. Эти методы, а также правила и специфика их применения будут рассмотрены в последующих параграфах настоящего учебника.

#### **IV этап. Разработка программы и проведение научного эксперимента.**

Для достижения цели исследований в области психологии и проверки выдвинутых гипотез определяются, как правило, следующие этапы экспериментальной работы:

1. **Организационный этап:** определяется база исследования (учреждение, предприятие и т.п.) и формируются экспериментальные (одна или несколько, в зависимости от цели и методов исследования) выборки.

2. **Методический этап:** определяется график и последовательность применения эмпирических методов исследования, производится подготовка инструментария.

3. **Практический этап:** в соответствии с методами исследования производится сбор эмпирических данных.

4. **Этап первичной обработки результатов исследования:** обработка, оцифровка и первичный анализ данных.

5. **Статистический этап:** полученные данные подвергаются статистическому анализу.

6. **Аналитический этап.** Полученные результаты анализируются с позиций методологической основы исследования, делаются выводы о достоверности гипотезы исследования.

## **V этап. Анализ результатов научного исследования.**

На данном этапе вносятся корректировки в концепцию и исследовательскую модель, формулируются новые гипотезы.

В работе Б.М.Теплова «О культуре научного исследования» [93] сформулированы следующие положения об использовании результатов исследований и о выводах.

1. Основным требованием исследователя-психолога к себе должно быть требование доказательности. Метод примеров особенно опасен, он не может быть использован в качестве метода доказательства. Однако пример законен и важен в области т.н. разведывательных (поисковых) работ, в частности для указания на то, что так бывает, так может быть.

2. Законность описательных работ. Не всякая работа открывает закономерности. Наука в целом не может быть описательной, а отдельное исследование может быть таким. Обязательным является не объяснение (особенно если оно вымучено, придумано, представляет собой псевдозаконмерности), а включение фактов, добытых исследованием, в контекст науки.

3. Вывод из работы. Вопрос о выводах из работы часто понимается чрезмерно широко, что выражается в стремлении сделать побольше выводов. В этом смысле стоит вспомнить слова П. П. Блонского о том, что больше внушает доверия та работа, в которой есть всего один вывод, но действительно достойный этого названия. Особенно это недопустимо, когда в выводы попадают положения заранее известные, трюизмы. Чрезвычайно опасен шаблон – в работах разных жанров могут быть очень разные выводы. Например, вывод может просто гласить: «описано то-то». Надо думать, что если стараются делать непосредственные выводы теоретического значения из любого частного исследования, то это нельзя квалифицировать иначе как «детскую болезнь» науки. Ведь ценность работы прежде всего решается включением в «историю вопроса», в систему фактов, ведущих к решению данной проблемы. А это возможно лишь в результате ряда работ, так как серьезные научные



проблемы решаются в итоге работ научных коллективов, а не отдельных исследований.

4. Количественная обработка материала. Конечно, психологи не пойдут по пути, предполагающему, что статистические операции могут придать смысл и содержание результатам, которые для самого исследователя не имеют ясного смысла и содержания. Но нужно подумать о культуре этого дела.

В заключение параграфа приведём слова выдающегося философа П.Фейерабенда: «Теория, выдвигаемая учёным, зависит не только от фактов, имеющихся в его распоряжении, но и от традиции, представителем которой он является, от математического аппарата, которым случайно владеет, от его вкусов, его эстетических взглядов, от мнения его друзей и других элементов, которые существуют не в фактах, а в мышлении теоретика и, следовательно, носят субъективный характер» [99].

### **Контрольные вопросы и задания**

- 1. Как связаны между собой следующие понятия: методологическая основа исследования, методы исследования, методики исследования?*
- 2. На какие группы подразделяются методы исследования по степени общности?*
- 3. Перечислите этапы научного исследования в психологии.*

## **1.2. Психодиагностический метод измерения и оценки**

Измерением называется метод исследования, направленный на выявление качественных и количественных характеристик объекта с помощью средств измерения. Процедура установления взаимозначного соответствия между множеством объектов (состояний) и множеством чисел называется количественным измерением (далее под термином измерение будет пониматься количественное измерение). Процедура установления соответствия между множеством объектов (состояний) и множеством классов качественных



характеристик, являющаяся по сути качественным измерением, называется оценкой.

Итак, измерение (в количественном смысле) – это приписывание чисел объектам или событиям, осуществляемое по определенным правилам, результатом которого является информация, закодированная в числовой форме. В ходе психологических измерений числовые коды приписываются определенным свойствам психики человека. Числовое представление психических качеств позволяет выявлять скрытые взаимосвязи между ними, что без математического аппарата было бы весьма затруднительно. Вместе с тем, как отмечает Г.В.Суходольский [92, с. 27], при измерении в психологии **«операции с числами должны адекватно отображать действия с вещами и обосновываться известными нам свойствами самих вещей, иначе получатся фиктивные результаты».**

Следовательно, ключевым вопросом психологического измерения является технология (инструмент) измерения интересующих исследователя психологических характеристик.

**Область психологической науки, разрабатывающая инструменты измерения и оценки индивидуально-психологических особенностей личности, называется психологической диагностикой.**

Термин «психодиагностика» ввёл Герман Роршах в 1921 г. Этим термином он назвал процесс обследования с помощью созданной им методики, и долгое время понятие психодиагностики относилось исключительно к методике Роршаха и другим проективным тестам. К концу XX века большинство западных психологических словарей трактуют психодиагностику исключительно для обозначения каких-либо психических расстройств.

В нашей стране направление исследований, получившее название «психологическая диагностика» появилось в 1960-х г.г. (термин «психодиагностика» был введён Б.Г.Ананьевым). Цели этих исследований охватывали достаточно широкий спектр изучения психологических и психофизиологических особенностей человека.

По определению Л.Ф.Бурлачука психодиагностический метод – это исследовательский метод психологической науки, характеризующийся измерительно-испытательной направленностью, за счёт которой достигается количественная и качественная оценка изучаемого явления. Психодиагностический метод конкретизируется в трёх основных диагностических подходах [10, с.104]:

**Объективный подход** – диагностика осуществляется на основе успешности (результативности) и способа (особенностей) выполнения деятельности.

**Субъективный подход** – диагностика осуществляется на основе сведений, сообщаемых испытуемым о себе, самоописания особенностей личности, поведения в тех или иных ситуациях.

**Проективный подход** – диагностика осуществляется на основе анализа особенностей взаимодействия с внешне нейтральным, как бы безличным материалом, становящимся в силу его неопределённости и слабоструктурированности объектом проекции.

По определению А.Г.Шмелёва [106], психодиагностика – это мост между наукой и практикой (наукой об индивидуальных психологических различиях (дифференциальная психология) и практикой постановки психологического диагноза). Теоретические основы психодиагностики задаются соответствующими областями психологической науки (общая психология, дифференциальная психология, возрастная психология, медицинская психология и т.д.). С другой стороны, психодиагностика – не только воплощение понятий соответствующих дисциплин в конкретных методиках, но и способ проверки истинности теоретико-психологических построений.

Успехи психодиагностики (т.е. успехи измерения и оценки) определяются, прежде всего, степенью разработанности теории измеряемого явления. Например, для получения эффективных интеллектуальных тестов потребовалась разработка научно обоснованных представлений о природе интеллекта. В настоящее время существует множество различных концепций интеллекта, в рамках которых разработаны различные психодиагностические методики.

Этим многообразием исходных концепций, кстати, и объясняются возможные несоответствия по результатам применения различных тестов: результаты каждой психодиагностической методики должны интерпретироваться с позиций концепции, лежащей в основе этой методики.

Современная психодиагностика может быть разделена на несколько взаимосвязанных частей, у каждой из которых свой предмет изучения:

1) методологические, теоретические и конкретно-методические принципы построения психодиагностических инструментов;

2) обоснование требований, которым должны удовлетворять измерительные психодиагностические методики (дифференциальная психометрика – измерительный фундамент психодиагностики);

3) выработка правил проведения обследования, способов обработки и интерпретации результатов;

4) особенности психодиагностики способностей, мотивов, черт личности, межличностных отношений, а также специфика применения соответствующих методик в различных областях жизнедеятельности людей (образование, профессиональная деятельность, семейные отношения, медицина, спорт и т.д.)

На практике полученные с помощью психодиагностической методики результаты либо соотносятся с определённой системой отсчёта (например, с тестовыми нормами, полученными в результате процедуры стандартизации методики, см. параграф 3.2), либо сравниваются между собой. В связи с этим можно говорить о двух типах диагноза:

1.«Диагноз клинический» [37]. Диагноз на основе констатации наличия или отсутствия какого-либо признака. В этом случае получаемые при диагностировании данные об индивидуальных особенностях психики испытуемого соотносятся с нормой или с каким-либо критерием.

2.«Диагноз статистический». Диагноз, определяющий особенности однородной (по возрасту, профессиональной принадлежности и

т.д.) выборки испытуемых по выраженности тех или иных психологических характеристик по сравнению с контрольной группой (выборкой стандартизации) или с другой однородной выборкой.

По Л.С.Выготскому [12] существуют следующие уровни психологического диагноза:

**1 уровень** – симптоматический, он ограничивается констатацией определённых особенностей или симптомов, на основании которых непосредственно строятся практические выводы;

**2 уровень** – этиологический диагноз, в ходе которого анализируются также причины возникновения симптомов;

**3 (высший) уровень** – типологический диагноз – определяется место и значение полученных данных в целостной динамической картине личности, с учетом сложной структуры личности.

Таким образом, психологическая диагностика является одним из основных эмпирических методов психологического исследования.

В зависимости от целей психодиагностической работы судьба поставленного психологом диагноза может быть различной. От этого существенно зависит содержание психологического заключения, его формулировки, отражающие суть полученных результатов (рис. 1.2).



**Рис. 1.2. Судьба психологического диагноза**

1. Диагноз, который передаётся клиенту (или его родителям, если клиентом является ребёнок до 12 лет). Заключение в этом случае пишется на языке, понятном клиенту. Сначала описываются положительные, ресурсные качества личности испытуемого. Отри-

цательные или недостаточно развитые характеристики сопровождаются конкретными рекомендациями по развитию или коррекции этих качеств.

2. Коррекционно-развивающая, консультационная или психотерапевтическая работа с испытуемым проводится специалистом-психологом (тем, который проводил обследование, либо его коллегой). В этом случае в заключении допускается специальная терминология, графическое представление результатов и т.д.

3. Диагноз передаётся другому специалисту (учителю, врачу и др.), который использует его с целью повышения эффективности своей работы с обследуемым. Заключение по результатам обследования в этом случае пишется с позиций и в терминологии профессиональной деятельности «заказчика».

При каждом из перечисленных вариантов передачи диагноза должны быть соблюдены профессионально-этические нормы психодиагностики, включающие в себя наиболее общие правила психодиагностического обследования, предупреждающие неоправданное или некомпетентное использование психодиагностических методик, а также являются кодексом профессиональной этики психолога, занимающегося диагностическими исследованиями.

Этические принципы деятельности психолога-диагноста в известной мере воспроизводят общечеловеческие нормы морали (профессиональный долг, соблюдение прав человека, обязательства по отношению к обществу, гуманизм и т.д.) Наряду с этими общечеловеческими принципами существует и ряд специфических принципов, вытекающих из особенностей взаимодействия «психолог-обследуемый». Психолог-диагност выступает в качестве эксперта личности обследуемого, его профессиональная деятельность нередко затрагивает наиболее интимные стороны жизни другого человека, содержание его сложных межличностных отношений, конфликтов. Объектом изучения в ряде случаев являются болезненные проявления личности, отклонения в поведении, последствия тяжёлых соматических или психических заболеваний. В таких, достаточно типичных для психолога ситуациях значение норм

профессиональной этике в психодиагностике сравнимо лишь с их значимостью в области медицинской деятельности.

Самое важное при проведении обследования и сообщении результатов испытуемым – это соблюдение профессионально-этических норм психодиагностики. Эти нормы включают в себя наиболее общие правила психодиагностического обследования, предупреждающие неоправданное или некомпетентное использование психологических тестов, а также являются кодексом профессиональной этики психолога, занимающегося диагностическими исследованиями. В случае нарушения психологом профессиональной этики, нанесения морального ущерба, любые позитивные результаты сводятся к нулю, подрывается авторитет коллег и психологической науки в целом.

**Принцип конфиденциальности** предполагает неразглашение результатов обследования без персонального соглашения на это испытуемого. Если речь идёт о несовершеннолетних, то на разглашение результатов требуется согласие родителей или заменяющих их лиц, несущих моральную и юридическую ответственность за детей. В тех случаях, когда психодиагностические процедуры проводятся в научных целях как часть экспериментального исследования, в отчёте имена и фамилии испытуемых шифруются (например, можно указать инициалы, либо порядковый номер).

**Принцип научной обоснованности** психодиагностической методики требует того, чтобы при обследовании использовались надёжные и валидные тесты, которым можно доверять. Очень важно помнить, что даже малейшее отклонение от инструкции и условий проведения тестирования может привести к серьёзным искажениям результата.

**Принцип соблюдения интересов испытуемого** требует безусловного уважения личности испытуемого, создания условий, при которых невозможна психическая травма. Каждый испытуемый имеет право получить от психолога информацию о результатах психодиагностического обследования. Эта информация должна



быть сообщена с неукоснительным соблюдением принципа конфиденциальности и принципа эффективности рекомендаций.

**Принцип ненанесения ущерба** предполагает, что результаты психодиагностического обследования нельзя использовать во вред тому человеку, который подвергается обследованию. В тех ситуациях, когда предполагается передача результатов обследования другому специалисту, то, во-первых, испытуемые должны быть проинформированы о направлениях диагностики и о том, кому и в каком виде эти результаты будут сообщены, а, во-вторых, специалисту-непсихологу следует передавать не сырые результаты тестирования, а специфические для его деятельности практические рекомендации, характеристики и т.п., выработанные по результатам психодиагностических процедур. Заметим также, что каждый испытуемый имеет право отказаться от обследования.

**Принцип объективности выводов** из результатов тестирования требует, чтобы они были научно обоснованными (т.е. объективно вытекали из результатов тестирования, проведённого при помощи валидных и надёжных методик), а не определялись и никак не зависели от субъективных установок тех, кто проводит тестирование или пользуется его итогами. При интерпретации результатов психодиагностического обследования необходимо учитывать ситуацию, в которой проводилось тестирование индивидуальные особенности испытуемых, их состояние и другие факторы.

**Принцип эффективности предлагаемых рекомендаций** предполагает, что такие рекомендации обязательно должны быть полезными для того человека, которому даются. Не разрешается предлагать человеку такие практические рекомендации, которые для него бесполезны или могут привести к нежелательным, непредсказуемым последствиям.

К специалистам, проводящим психодиагностическое обследование, предъявляются следующие квалификационные требования: хорошая теоретическая подготовка, знание методологических основ применяемых психодиагностических методик; доскональное

знание применяемых методик, правил и особенностей работы с ними.

В Приложении 1 представлен Этический кодекс и Клятва российского психолога, принятые V съездом Российского психологического общества 14 февраля 2012 г.

Рассмотрим далее специфику различных терминов, имеющих отношение к методам психологической оценки.

В современной западной психологической науке существует термин «психологическое тестирование». Этим термином обозначается всё связанное с разработкой и применением любых психологических тестов, как инструментов измерения индивидуальных различий.

В настоящее время всё шире используется термин «психологическая оценка» – более широкое понятие, чем «психологическое тестирование». «Оценка» – это сбор и интеграция данных, которые могут быть получены не только с помощью тестов (но и с помощью беседы, интервью, наблюдения, т.е. методов внетестовой диагностики).

В соответствии с определением в словаре С.И.Ожегова оценка – это мнение о ценности, уровне или значении кого- или чего-нибудь. Термин «оценка» имеет многолетние традиции использования в педагогике. Так, в педагогическом словаре [32, с.231] оценка педагогическая определяется как процесс соотнесения результата деятельности или поведения воспитанника или хода самой деятельности с заранее заданными эталонами. Б.Г.Ананьевым введён также термин «парциальная оценка» – отдельные оценочные обращения и воздействия педагога на воспитанника, не представляющие собой квалификацию его успешности, а являющиеся лишь эмоциональной реакцией педагога (одобрение, замечание, порицание, поддержка, сарказм, упрёк, нотации и т.п.). Парциальные оценки оказывают также «последствие», влияют на отношения между участниками образовательного процесса, на самооценку обучающихся, формирование интересов и т.д.



В заключение параграфа отметим, что цели использования методов измерения в психологии могут быть разделены на практические и научные. Это разделение носит условный характер, поскольку за каждым сугубо практическим использованием любого средства (методики) измерения (прибора, теста, анкеты) стоит концепция разработчика. Поэтому для грамотного применения методов психологической оценки в практической деятельности необходима теоретическая подготовка, знание методологических основ применяемых методов и методик, правил и особенностей работы с ними.

### **Контрольные вопросы и задания**

1. *Дайте определения понятиям «психологическая диагностика», «психологическое тестирование» и «психологическая оценка».*
2. *Дайте определение понятиям «измерение» и «оценка».*
3. *Выпишите определения термина «психодиагностика» у различных авторов и произведите их сравнительный анализ.*
4. *Перечислите профессионально-этические принципы проведения психодиагностического обследования.*
5. *Проведите сравнительный анализ кодексов профессиональной этики Российского психологического общества (см. Приложение 1) и Американского психологического общества (см. [4]).*

### **1.3. Классификация методов психологической оценки**

Методы психологической оценки делятся на две группы: методы высокого уровня формализации и малоформализованные методы (рис. 1.3).

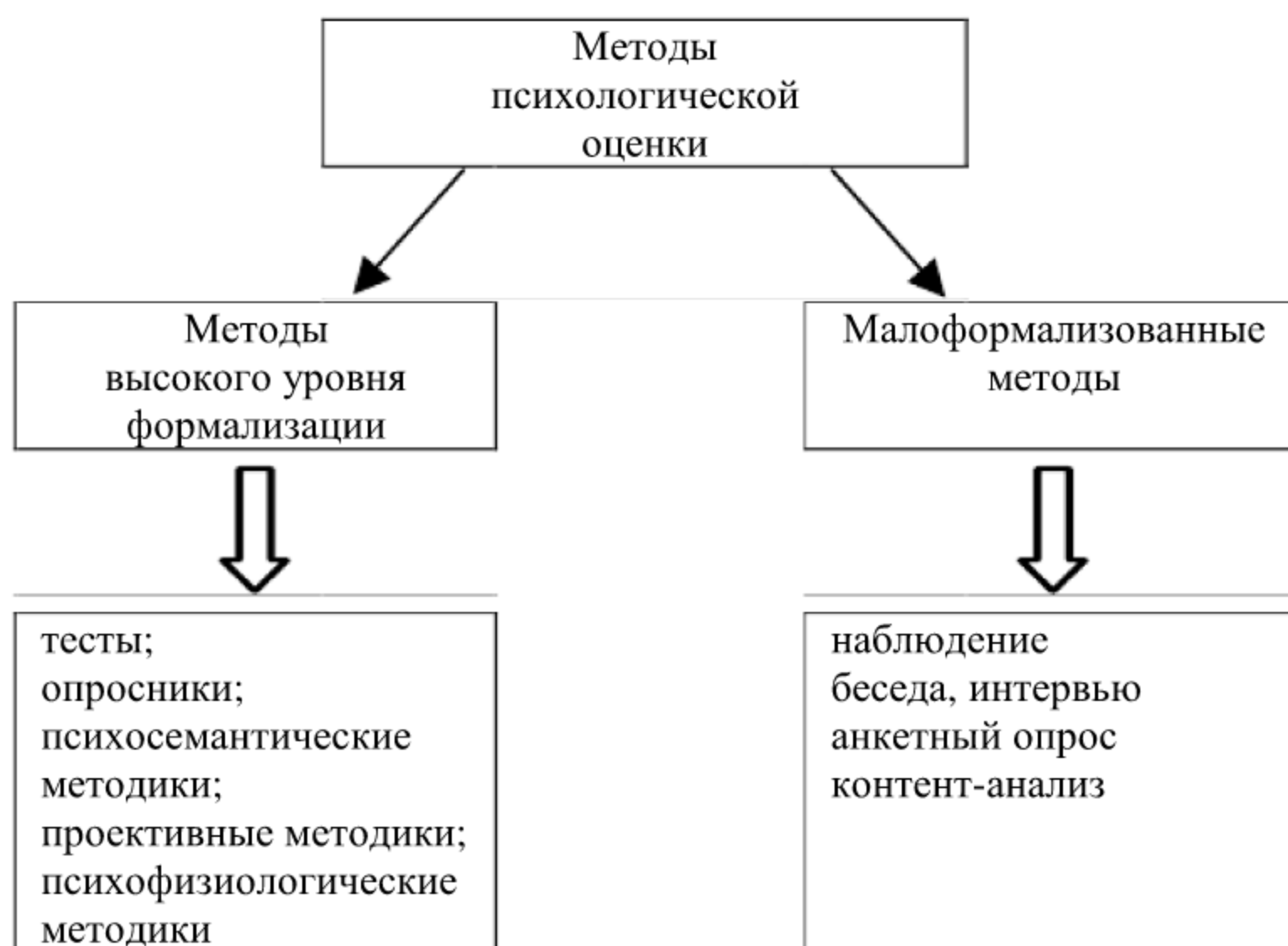
К методам высокого уровня формализации относятся следующие типы методик: тесты; опросники; психосемантические методики; методики проективной техники; психофизиологические методики. Эти методики называются строгоформализованными.

Для строгоформализованных методик характерны:

- **стандартизация процедуры проведения** (регламентация): точное соблюдение инструкций, строго определённые способы

предъявления стимульного материала, невмешательство исследователя в деятельность испытуемого;

- **стандартизация оценки выполнения:** единообразие проведения обработки и представления результатов диагностических обследований;
- **надёжность:** точность измерения, а также устойчивость результатов к действию посторонних случайных факторов;
- **валидность:** «понятие, указывающее, что тест измеряет и насколько хорошо он это делает» (А.Анастаси). В данном, ставшем классическим, определении под словом «тест» подразумевается термин «строгифомализованная методика». Валидность – это комплекс сведений о том, относительно каких групп психологических свойств личности могут быть сделаны выводы с помощью данной методики, а также о степени обоснованности выводов при использовании конкретных тестовых оценок или других форм оценивания. Иными словами, валидность – это определённая гарантия соответствия результатов, получаемых при помощи данной методики, замыслу её создателя.



### **Рис. 1.3. Классификация методов психологической оценки**

Строгоформализованные методики позволяют собрать диагностическую информацию в относительно короткие сроки и в таком виде, который даёт возможность количественно и качественно сравнивать индивидов между собой.

К малоформализованным методам психологической оценки относятся: наблюдение, беседа и интервью, анкетный опрос и анализ продуктов деятельности (контент-анализ). Применение этих методов даёт очень ценные сведения об испытуемом, но процедуры достаточно трудоёмки (наблюдения за обследуемым могут осуществляться в течение нескольких месяцев) и в большей степени основаны на профессиональном опыте, психологической подготовленности специалиста, производящего психологическое оценивание. Только высокий уровень культуры проведения психологических наблюдений, бесед помогает избежать влияния случайных и побочных факторов на результаты обследования или испытания.

При проведении психологической оценки строгоформализованный и малоформализованный методы, как правило, взаимно дополняют друг друга. В полноценном обследовании необходимо их гармоничное сочетание. Например, сбору данных с помощью тестов должен предшествовать период ознакомления с обследуемым с помощью интервью, беседы, наблюдения.

### **Контрольные вопросы и задания**

- 1. На какие группы делятся методы психологической оценки? Охарактеризуйте каждую группу.*
- 2. Перечислите требования, предъявляемые к методикам, разработанным в рамках строгоформализованного метода психологической оценки.*

## 1.4. Шкалы измерения и экспериментальная выборка

Существует четыре типа шкал измерений (и, соответственно, четыре типа переменных): номинальная, порядковая, интервальная, относительная (шкала равных отношений).

*Номинальная (номинативная) шкала* служит исключительно для качественной классификации переменных, т.е. переменные могут быть измерены лишь в терминах принадлежности к некоторым существенно различным, непересекающимся группам (классификационным ячейкам). Примером номинальной шкалы является классификация людей по типам темперамента (сангвиники, холерики, флегматики, меланхолики). Один из видов номинальной шкалы – дихотомическая шкала, в которой признак может принимать только два значения (например, пол испытуемых).

При измерении признака по номинативной шкале, можно подсчитать частоты встречаемости различных наименований в выборке, а затем работать с этими частотами с помощью математических методов. Например, сопоставить распределение респондентов по полу с равномерным распределением.

*Порядковая шкала* – это шкала измерения значений признака, которые могут быть упорядочены по степени наличия измеряемого психологического качества. Однако, в порядковой шкале неизвестно расстояние между различными значениями переменной, а известно лишь, что они образуют последовательность (по возрастанию или убыванию величины измеренного психологического качества). Примерами порядковой переменной могут служить как количественные значения, полученные испытуемыми в сырых баллах по любой нестандартизированной психодиагностической методике, так и качественные значения признака, которые можно упорядочить.

Математическая обработка данных, измеренных в порядковой шкале, в большинстве случаев основывается на ранжировании, при котором наименьшему значению переменной соответствует ранг 1,

следующему значению – ранг 2 и т.д., при этом одинаковым значениям переменной приписываются одинаковые ранги.

*Интервальная шкала* – это шкала значений переменной, позволяющая не только упорядочивать испытуемых, но и численно выражать и сравнивать различия между ними. Основное понятие этой шкалы – интервал – постоянная на всех участках шкалы доля измеряемого свойства между соседними позициями на шкале. Интервальное измерение – это такое присвоение чисел объектам или их свойствам, при котором равные разности чисел соответствуют равным разностям значений измеряемого признака. В психодиагностических измерениях равноинтервальными можно считать лишь шкалы в единицах стандартных показателей (Z- и T-показатели, стеноны, станайны).

Для обработки и анализа эмпирических данных, измеренных в шкале интервалов, допустимо подавляющее большинство методов статистической обработки данных, соответствующих цели и задачам исследования.

*Шкала равных отношений* отличается от интервальной шкалы тем, что имеет наличие абсолютной нулевой точки отсчета, указывающей на полное отсутствие измеряемого свойства. Шкала отношений является самой информативной шкалой, допускающей любые математические операции. Вместе с тем, при проведении статистической обработки данных, как правило, не делается тонкого различия между свойствами интервальной шкалы и шкалы отношений. В психологии шкала отношений используется в экспериментальных психофизических исследованиях, а также при измерении психофизиологических характеристик. В психодиагностике познавательной сферы и личностных особенностей измерения производятся в номинальной, порядковой и интервальной шкалах. Всегда можно перейти от более информативной шкалы – к менее информативной. Например, от интервальной шкалы значений IQ в баллах по шкале Векслера можно перейти к шкале порядка с градациями Н, НС, С, ВС, В (соответственно низкий уровень IQ, ниже среднего, норма, выше среднего, высокий). В некоторых случаях

такое понижение шкалы имеет смысл, например, позволяет упростить вычислительные процедуры, повысить наглядность результатов, но часть информации при этом безвозвратно утрачивается.

В заключение параграфа рассмотрим два понятия, имеющих непосредственное отношение к применению методов психологической оценки. Это понятия генеральной совокупности и выборки.

**Генеральной совокупностью** называется «множество объектов, эквивалентных по конечному множеству свойств» [21, с.304]. Иными словами, это множество людей, обладающих определенными характеристиками (возраст, пол, образование, профессия и т.п.), на которое распространяются результаты проведенного исследования. Если в психологическом эксперименте участвуют все представители генеральной совокупности, то такое исследование называется полным или сплошным; при этом полученная информация должна получиться исчерпывающей. На практике такие эксперименты не проводятся, так как исследовать генеральную совокупность, во-первых, в большинстве случаев нереально, а во-вторых, нецелесообразно. Для получения полноценной информации в соответствии с целью, гипотезой и задачами эксперимента достаточно обследовать выборку испытуемых, удовлетворяющую определенным требованиям по своему составу и объему.

**Выборкой** называется любая группа элементов (испытуемых, респондентов), выделенная из генеральной совокупности для проведения эксперимента. Выборка должна быть однородной и репрезентативной. Основаниями для формирования однородной выборки могут служить различные характеристики, такие как возраст, образование, социальное положение, профессия, национальность, наличие или отсутствие определенных заболеваний и т.д., в зависимости от целей исследования. Репрезентативной называется такая выборка, в которой все основные признаки генеральной совокупности представлены в той же пропорции, т.е. репрезентативная выборка представляет собой меньшую по размеру точную модель генеральной совокупности.



Чаще всего для формирования репрезентативной выборки используется метод стратифицированной случайной выборки. Для этого элементы генеральной совокупности разбиваются на страты (группы) в соответствии с некоторыми характеристиками и в дальнейшем производится случайная выборка из каждой страты. Так, при определении национальных норм для психодиагностических методик важными стратификационными переменными обычно являются возраст, пол и социальное положение, а также переменные, связанные со спецификой методики. Например, при стандартизации теста интеллекта Лордж-Торндайка основой дальнейшей стратификации групп, однородных по указанным выше признакам, послужили факторы, наиболее связанные, по мнению разработчиков, с уровнем интеллекта: процент грамотного взрослого населения, пропорция профессиональных рабочих в популяции, процент домовладельцев и среднее количество снимающих жилье. Каждая социальная группа была классифицирована по этим факторам на «очень высокий», «высокий», «средний», «низкий» и «очень низкий» уровни и для каждого из этих уровней в каждой группе были получены отдельные нормы [29, с.215].

Объем выборки зависит от целей исследования, однородности изучаемого явления и ограничений статистических методов обработки результатов. Как правило, при проведении психологических исследований используются средние по численности выборки (от 30 до 200 человек). При конструировании психодиагностических методик, исследовании их валидности, надежности и определении тестовых норм используются большие выборки (более 200 человек). Объем выборки стандартизации тестовой методики зависит от характеристик генеральной совокупности, на которую распространяются результаты исследования, и может быть очень большим. Так, для выработки норм для общей популяции детей школьного возраста необходима выборка объемом 10 000 испытуемых [29, с. 214]. Вместе с тем отметим, что репрезентативность выборки является более важной характеристикой, чем объем, поскольку именно состав выборки (её репрезентативность) влияет на правомочность

переноса полученных результатов на всю генеральную совокупность.

Объём выборки важен для применения статистического аппарата. В статистике различают малую ( $n < 30$ ), среднюю ( $30 < n < 200$ ) и большую выборку ( $n > 200$ ). Это деление носит условный характер и зависит от целей исследования. Статистическая значимость полученных результатов напрямую связана с объёмом выборки. На стадии планирования эксперимента следует соотносить объём выборки с ограничениями тех статистических критериев, которые предполагается применять для анализа результатов исследования. Например, если в ходе обработки данных исследования использовать факторный анализ, то объём выборки испытуемых должен быть не менее 100 человек.

Если в психологическом исследовании участвует несколько выборок испытуемых, то важной характеристикой, определяющей выбор метода статистической обработки результатов, является их связность. Выборки называются несвязными (независимыми), если процедура эксперимента и полученные результаты измерения некоторого свойства у испытуемых одной выборки не оказывают влияния на особенности протекания этого же эксперимента и результаты измерения этого же свойства у испытуемых другой выборки. В противном случае (когда это влияние есть) выборки называются связными или зависимыми. Так, при сопоставлении показателей, полученных на одной группе испытуемых в разных условиях измерения либо с течением времени, выборки считаются связными. Если же сопоставляются показатели, полученные по той же методике в двух группах испытуемых (например, в контрольной и экспериментальной), то эти выборки считаются несвязными.

### **Контрольные вопросы и задания**

1. *Охарактеризуйте типы шкал измерения.*
2. *Что такое генеральная совокупность и выборка?*
3. *Каковы требования, предъявляемые к выборке испытуемых?*



## **Глава 2.**

### **История развития методов оценивания в психологии и образовании**

#### **2.1. История развития методов психологической диагностики и оценки**

Интерес к индивидуальным различиям между людьми существует с древних времён. В Библии отражены способы диагностики по особенностям поведения (например, как отбирать воинов для битвы по тому, как человек пьёт воду из реки). Из истории древних цивилизаций известны различные способы диагностики индивидуальных различий. Так, в Древнем Китае была создана целая система отбора чиновников, в которой учитывались различные аспекты, от специальных способностей (счёт, письмо) до поведения в быту. Разного рода испытания практиковались и в Древней Греции. Так, выдающийся философ и математик Пифагор допускал в свою школу лишь тех, кто прошёл через ряд серьёзных экстремальных испытаний, в которых нужно было продемонстрировать достоинство (будучи выставленным на всеобщее осмеяние) и присутствие духа (например, ночью в страшной пещере). Пифагор также считал, что характер человека отражается в смехе и походке. Платон отмечал, что природа людей различна, поэтому можно сделать всё в большем количестве, лучше и легче, если выполнять одну какую-нибудь работу соответственно своим природным задаткам.

Различные типологии людей также известны с древних времён. Так, типология темпераментов Гиппократ (холерик, сангвиник, флегматик и меланхолик) используется до сих пор. Известна также типология характеров Теофраста (скупой, лгун, хвастун и т.д.), с помощью которой в зависимости от характерных черт конкретного человека его можно было отнести к той или иной категории и та-

ким образом прогнозировать его поведение. Во многих русских и европейских сказках претенденты на руку принцессы и царство должны были пройти через испытания, где требовалось проявить незаурядные интеллектуальные, личностные и физические способности.

Таким образом, человечество всегда характеризовалось стремлением к обнаружению и учёту индивидуальных различий (физических, физиологических, психологических), которые рассматривались как единое целое. Впоследствии, с развитием наук о человеке, составляющие индивидуальных различий разделились. Психологической наукой в настоящее время изучаются индивидуально-психологические различия между людьми, которые и являются предметом психодиагностики как науки.

«Диагноз» в переводе с греческого означает распознавание чего-либо (болезни в медицине, неисправности в работе автомобиля и т.д.) Психологическая диагностика предназначена для сбора информации об особенностях человеческой психики.

Психологическая диагностика, как самостоятельная область психологической науки, начала складываться на рубеже XX в., но индивидуальные различия между людьми оценивались и нашими далёкими предками.

Родоначальником научного изучения индивидуальных различий считается английский исследователь **Фрэнсис Гальтон** (1822 – 1911). Одарённость Фрэнсиса Гальтона (он был младшим, седьмым ребёнком в состоятельной и известной в научных кругах семьи) проявилась рано: с полутора лет знал все буквы алфавита, самостоятельно читал с двух с половиной лет, писал с трёх лет. С 1838 обучается медицине (Бирмингемский госпиталь, медицинская школа Лондона), в 1839 – в Королевском колледже на медицинском отделении. В 1840 поступил в Кэмбриджский университет для занятий математикой и естественными науками [26].

В 1849 Ф.Гальтон публикует первое научное сообщение, посвященное разработке телетайпа – печатного телеграфа. Он много путешествовал, в том числе с экспедициями по Ближнему Востоку и

Южной Африке. В 1854 получил золотую медаль Географического общества за отчёт об африканском путешествии. В 1855 публикует книги «Искусство путешествовать» и «Заметки о современной географии». С 1856 г. являлся членом Королевского общества. С конца 1850-х занимался климатологией и метеорологией, опубликовал работу о климате Занзибара, первым начал выпускать метеорологические карты Европы, открыл феномен антициклона.

После выхода книги Чарльза Дарвина «Происхождение видов» Фрэнсис Гальтон заинтересовался проблемой наследования различных признаков у человека и животных. В 1865 г. он публикует статью «Наследственный талант и характер», в 1869 г. выходит его книга «Наследственный гений», в которой анализируется обширный материал по «выдающимся людям», обсуждается феномен нормального распределения признаков в человеческой популяции.

В конце 1870-х годов Фрэнсис Гальтон разрабатывает методологию психометрических исследований. Публикует множество статей, изобретает первые приборы для психометрических опытов.

В 1883 г. Фрэнсис Гальтон разработал систему измерения психофизиологических качеств человека, предположив, что уровень интеллектуальной одарённости человека напрямую связан с параметрами продуктивности простейших психических процессов. Он был убеждён, что представители социальной элиты и биологически, и интеллектуально превосходят представителей социальных низов, а женщины гораздо менее талантливы и умны, чем мужчины. В 1884 году Гальтон организовал на Лондонской Международной выставке здоровья антропологическую лабораторию, где каждый посетитель, заплатив 3 пенса и заполнив анкету, мог проверить свои способности по 17 показателям и определить свои вес, рост, мышечную силу, остроту зрения и слуха, объём лёгких, различение цветов, запоминание объектов и т.д. Ф. Гальтон обследовал свыше 10 тысяч испытуемых. В результате оказалось, что по измеренным показателям деятели науки ничем особым не отличаются от рядовых «средних» посетителей выставки, а женщины по ряду показателей превосходили мужчин.

Несмотря на то, что Фрэнсису Гальтону не удалось создать инструмент измерения интеллектуальных способностей человека, его трудами были заложены научные основы психодиагностики, и, кроме того, им впервые был предложен термин «тест», как инструмента измерения индивидуальных различий. Для обработки результатов Гальтон впервые в психологии применил статистические процедуры, в 1888 году им на основе работ К. Пирсона был предложен метод вычисления коэффициента корреляции для обработки результатов психологических исследований. Фрэнсисом Гальтоном был также обоснован закон регрессии наследственных признаков, что внесло значительный вклад в развитие генетики.

Не менее значительным был вклад Ф.Гальтона в развитие криминалистики: в ходе психометрических исследований им была собрана коллекция отпечатков пальцев и доказано, что папиллярные линии образуют четыре основных типа рисунков. Математический расчет вероятности совпадения отпечатков 10 пальцев, сделанный Гальтоном, показал, что совпадение отпечатков пальцев двух людей невозможно. Фрэнсис Гальтон разработал формулы для идентификации отпечатка, в которых учитывались мельчайшие детали: угол наклона папиллярной линии, коэффициенты ее закругления, форма треугольника, который составляют линии. В 1892 году Гальтон опубликовал в Лондоне книгу «Finger prints», в которой изложил результаты своих исследований. В 1895 году дактилоскопия была взята на вооружение Скотланд-Ярдом, и с 1900 года дактилоскопический метод Гальтона стал основным методом идентификация преступников.

В последние годы жизни ученый занимался разработкой основных положений науки евгеники о создании идеального во всех отношениях человека. Фрэнсис Гальтон был первым, кто поставил вопрос об измерении личностных особенностей. Он считал, что точное измерение характера даёт «статистика поведения каждого человека в малых ежедневных делах».

Ф. Гальтону принадлежит также идея применения методик проективной техники для диагностики личности. Им был предложен

метод «ассоциации идей» (изучение ассоциативных реакций испытуемого на различные стимулы).

Своими исследованиями Ф. Гальтон положил основу создания школы изучения интеллекта, у него появились ученики и последователи, одним из которых являлся **Джеймс Мак-Кин Кеттелл** (1860-1944), обосновавший в своих исследованиях пути преобразования психологии в точную науку.

В ходе проведённых исследований Д. Кеттеллом из списка использовавшихся к тому времени 50 методов экспериментально-психологических измерений было отобрано 10, включавших различного рода измерения чувствительности, времени реакции, количества звуков, воспроизводимых после однократного прослушивания и т.п. Он считал, что эти тесты, являющиеся в основном тестами элементарных психических функций, наиболее подходящими для измерения ума. Но в отличие от экспериментально-психологического исследования, смысл применения этих тестов был принципиально иным. Задачей психологического эксперимента является выяснение зависимости психического акта от внешних и внутренних факторов (например, характера восприятия от внешних раздражителей). При тестировании же психолог регистрирует индивидуальные различия психических актов, не изменяя внешних условий их осуществления.

В своей статье «Умственные тесты и измерения» («Mental Test and Measurement»), опубликованной в 1890 году в журнале «Mind» Кеттелл писал, что психология не сможет стать такой же обоснованной и точной, как физические науки, если будет обходиться без эксперимента и измерения. Первый шаг в этом направлении может быть сделан с помощью применения серии умственных тестов к значительному количеству индивидов. Результаты могут иметь значительную научную ценность в открытии постоянства психических процессов, их взаимозависимости и изменений в различных обстоятельствах. Во-вторых, статистический анализ результатов тестирования является залогом становления новой психологии, психологии, опирающейся на точное знание. Тем самым



Д. Кеттеллом была провозглашена необходимость стандартизации тестов для того, чтобы стало возможным сравнение результатов, полученных разными исследователями на разных испытуемых.

Таким образом, Дж.Кеттеллом были сформулированы **два концептуальных положения**, которые положили начало психодиагностике как науки:

1. Применение серии тестов к большому количеству индивидов позволит открыть закономерности психического развития индивидов и тем самым превратит психологию в точную науку.

2. Научная и практическая ценность тестов возрастёт, если условия их проведения будут однообразными.

Вместе с тем, содержание известных к тому времени тестов не отличалось от лабораторных исследований экспериментальной психологии. Фрэнсис Гальтон, написавший послесловие к статье Дж.Кеттелла, высказал ряд критических замечаний и фактически поставил вопрос о валидности методов изучения умственных способностей. Низкую валидность методов измерения интеллектуальных способностей, которые применяли Кеттелл и Гальтон в своих исследованиях, подтвердили и результаты применения этих тестов в области образования. Так, Ф. Боаз (США) в 1891 г. протестировал около 1500 школьников и практически не обнаружил соответствия между своими данными и теми данными, которые представили учителя, не обращавшиеся к помощи интеллектуальных тестов Дж.Кеттелла. А опубликованные в 1901 г. результаты исследований ученика Кеттелла К.Висслера, показали полную непригодность существующих к тому времени тестов для измерения интеллектуальных способностей. Так, коэффициенты корреляции академической успеваемости школьников с результатами тестирования были следующими: 0,02 – со временем реакции; -0,02 – с распознаванием цветов; -0,08 – с показателями динамометрии. Эти результаты свидетельствовали о необходимости иного подхода к оценке интеллектуальных способностей, а также о необходимости доказательства пригодности каждого нового метода для этих измерений.



Эта задача была успешно решена французским психологом, доктором медицины и права, **Альфредом Бине** (1857 – 1911). Он проводил научные исследования в русле требований практики об измерении высших психических функций, обеспечивающих приобретение знаний и успешное выполнение сложной приспособительной деятельности. С исследований Бине, его коллег и последователей началась дифференциация тестов интеллекта, и тестов, измерявших индивидуальные различия, не связанные с интеллектом (мышечная сила, объём лёгких и т.п.)

В 1904 г. Министерство образования Франции поручило Бине заняться разработкой методик, с помощью которых можно было бы отделить детей, способных к учению, но ленивых и не желающих учиться, от детей, страдающих врождёнными дефектами и не способных учиться в нормальной школе.

А. Бине предложил включать в содержание тестов такие задания, о которых известно, в каком возрасте нормальные дети их правильно выполняют. Показателем интеллекта в шкалах Бине был умственный возраст, который определялся по успешности выполнения тестовых заданий (возраст, задания для которого успешно выполнял ребёнок). Несовпадение умственного и хронологического возраста считался показателем либо умственной отсталости, либо одарённости. Для каждого возраста детей А.Бине разработал серию заданий. Приведём примеры заданий теста Бине для детей 7 и 8 лет.

*Для детей 7 лет:*

1. Различать правую и левую стороны.
2. Описать предложенную картинку.
3. Выполнить несколько поручений.
4. Назвать общую стоимость нескольких монет.
5. Назвать показанные четыре основных цвета.

*Для детей 8 лет:*

1. Сравнение двух объектов по памяти. Установление сходства между ними.
2. Обратный счет от 20 до 1.

3. Обнаружение пропусков в изображениях людей (4 задания).
4. Назвать день, число, месяц, год.
5. Повторить ряд из пяти однозначных чисел.

Л.С.Выготским был высказан ряд критических замечаний по содержанию заданий теста Бине, его практической значимости, но, вместе с тем, он высоко оценил идею и методический принцип, заложенный в основу теста.

Тест Бине претерпел существенные изменения после работы по его проверке и стандартизации, проведённой в Стэнфордском университете под руководством Левиса Мэдисона Термена. Этот вариант тестовой шкалы Бине был назван шкалой Стэнфорд-Бине.

В ходе разработки в базовую методику было введено большое количество изменений. По сравнению со шкалой Бине было добавлено более трети новых заданий, ряд старых был либо переделан, либо отброшен или переадресован другим возрастным группам. Фактически уже первая редакция методики представляла собой новый тест. В дальнейшем тест неоднократно радикально совершенствовался [11].

Шкала Стэнфорд-Бине включает задания, направленные на исследование широкого диапазона способностей – от простого манипулирования до абстрактного рассуждения. На ранних возрастных уровнях тесты требуют главным образом зрительно-моторной координации, перцептивного различия, способности понимать инструкцию (в заданиях типа складывания кубиков, нанизывания бус, подбора геометрических фигур), а также способности узнавать предметы, предъявляемые в виде игрушечных моделей или изображений на карточках. На высших возрастных уровнях в наибольшей степени представлены тесты, использующие вербальное содержание заданий. Среди них словарный тест (объяснение значения слов), аналогии, завершение предложений, определение абстрактных понятий, интерпретация пословиц. Некоторые тесты направлены на характеристику степени плавности и беглости речи (быстрое называние не связанных между собой слов, подбор рифм, построение предложений с заданными словами). Среди заданий

батарей широко представлены тесты общей осведомленности, знания норм общественной жизни, правил поведения (ответы на вопросы, интерпретация ситуаций, обнаружение несоответствий в сюжетных картинках или рассказах). В шкалу включен ряд тестов памяти, пространственной ориентации (визуальная репродукция фигур, лабиринты, складывание и разрезание бумажных объектов и т. д.). На более высоких возрастных уровнях анализируется степень усвоения некоторых навыков, приобретенных в школе (умение читать, знание арифметики).

Методика допускает возможность получения широкой качественной информации о методах работы испытуемого, способах решения им задач. Большие возможности предоставляются и наблюдению за личностными качествами: уровнем активности и мотивации, уверенности, настойчивости, сосредоточенности и т. д. Сложная процедура проведения обследования и интерпретации полученных результатов, необходимость строгого соблюдения стандартов требуют высокой квалификации экспериментатора.

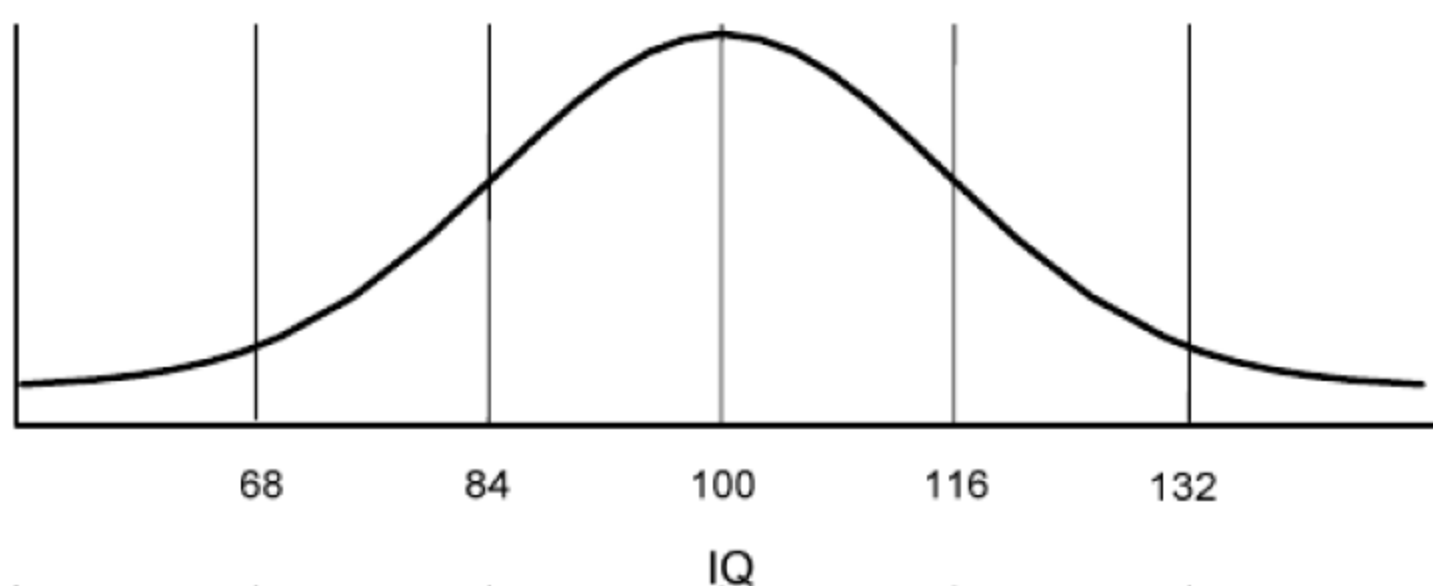
По применению шкалы Стэнфорд-Бине накоплен огромный опыт, включающий фактические данные и их интерпретацию. По широте использования эта методика занимает в зарубежной психодиагностике одно из ведущих мест среди тестов интеллекта. Длительность использования и широта распространения сделали систему отсчета оценок шкалы Стэнфорд-Бине стандартом для других психометрических тестов. Распределение результатов IQ-показателей шкал Стэнфорд-Бине положено в основу классификации степеней умственной отсталости, широко распространенной в зарубежной психодиагностике. Таким образом, по сравнению с тестами Бине, в шкале Стэнфорд-Бине было два основных нововведения:

- 1) введение в качестве показателя по тесту коэффициента интеллектуальности IQ (Intellectual Quotient), выводящегося из отношения между умственным возрастом (МА – mental age) и хронологическим возрастом (СА – chronological age), умноженного на 100:

$$IQ = (MA / CA) * 100 \quad (2.1)$$

2) применение статистического критерия оценки тестирования, для чего было введено понятие статистической нормы. Норма стала тем критерием, с которым производится сравнение индивидуальных тестовых показателей и на его основе – психологическая интерпретация результатов измерения.

Шкала Стэнфорд-Бине рассчитана на детей в возрасте от 2.5 до 18 лет. Она состоит из заданий разной трудности, сгруппированных по возрастным критериям. Для каждого возраста среднее значение шкалы  $\mu = 100$ , а статистическая мера рассеяния, стандартное отклонение от среднего  $\sigma = 16$  (рис. 2.1).



**Рис. 2.1 Распределение оценок в шкале Стэнфорд-Бине**

Все индивидуальные показатели по тесту, попадавшие в интервал  $[\mu - \sigma, \mu + \sigma]$ , т.е. ограниченные числами 84 и 116, считались нормальными, соответствующими возрастной норме. Если тестовый показатель был выше тестовой нормы (больше 116), ребёнок считался одарённым, а если ниже 84 – то умственно отсталым.

Шкала Стэнфорд-Бине получила популярность во всём мире. Она имела несколько редакций (1937, 60, 72, 86 г.г.) В последней редакции она применяется и в настоящее время. Показатель IQ, получаемый по шкале Стэнфорд-Бине, на долгие годы стал синонимом интеллекта, а сама методика – эталонной. Вновь создаваемые интеллектуальные тесты стали проверяться путём сопоставления с результатами шкалы Стэнфорд-Бине.

Следующий этап развития психологического тестирования характеризуется изменением формы проведения тестового испытания. Все тесты, созданные в начале XX века, были индивидуальными и позволяли вести опыт только с одним испытуемым, что существенно ограничивало их распространение и не отражало потребности практики диагностировать большие массы людей с целью отбора наиболее психологически подготовленных к тому или иному виду деятельности. Так, в период первой мировой войны появилась необходимость как можно быстрее отобрать и распределить полутора миллионную армию рекрутов по различного рода армейским службам, школам и училищам. В связи с этой необходимостью в США были созданы армейские тесты для группового тестирования людей. Форма «Альфа» предназначалась для людей, знающих английский язык, «Бета» – для неграмотных и иностранцев. После окончания войны применение групповых методик получило широкое распространение в системе образования, промышленности и в армии.

На протяжении первой половины XX века было создано множество разнообразных тестов, которые тщательным образом стандартизировались на больших выборках, отличались высокой надёжностью и хорошей валидностью. Вместе с тем, были выявлены ограниченные возможности комплексных тестов интеллекта из-за ограничений на общий объём методики (стремление «объять необъятное»). Кроме того, успешность выполнения различных частей теста у людей была различной. Требовалась дополнительная и дифференцированная информация об особенностях психики человека. В результате возникла потребность в разработке тестов специальных способностей и тестов достижений.

Толчком для развития тестов специальных способностей стало мощное развитие профессионального консультирования, а также профессионального отбора и распределения персонала в промышленности и военном деле. Стали появляться тесты механических, музыкальных, канцелярских и других способностей. Каждый из



таких тестов диагностировал лишь одну узкую специальную способность.

Особое направление в психологической диагностике связано с разработкой различных методов диагностики личности. К их числу относятся опросники и проективные методики. Прототипом личностных опросников был опросник, разработанный в 1919 г. американским психологом Робертом Вудвортсом. Он назывался Бланк данных о личности и был предназначен для выявления и отсеивания с военной службы лиц с невротической симптоматикой.

Другим известным методом диагностики личности являются проективные методики, родоначальником которых считается метод словесных ассоциаций (эксперименты Френсиса Гальтона, в которых он предлагал испытуемым отвечать на слово-раздражитель первой пришедшей в голову словесной ассоциацией, при этом фиксировалось время ответа). Позднее эта методика получила развитие в исследованиях многих психологов и психиатров.

Одна из самых популярных проективных методик была разработана в 1921 г. швейцарским психиатром Германом Роршахом, который в ходе обследования предъявлял испытуемым – психически больным – чернильные пятна (рис. 2.2).



**Рис. 2.2. Образец одного из заданий теста Г.Роршаха**

Г.Роршахом был замечен тот факт, что люди со схожими психическими расстройствами склонны интерпретировать чернильные пятна схожим образом. В результате те характеристики ответов, которые можно было соотнести с различными психическими заболеваниями, постепенно объединялись в систему показателей пси-



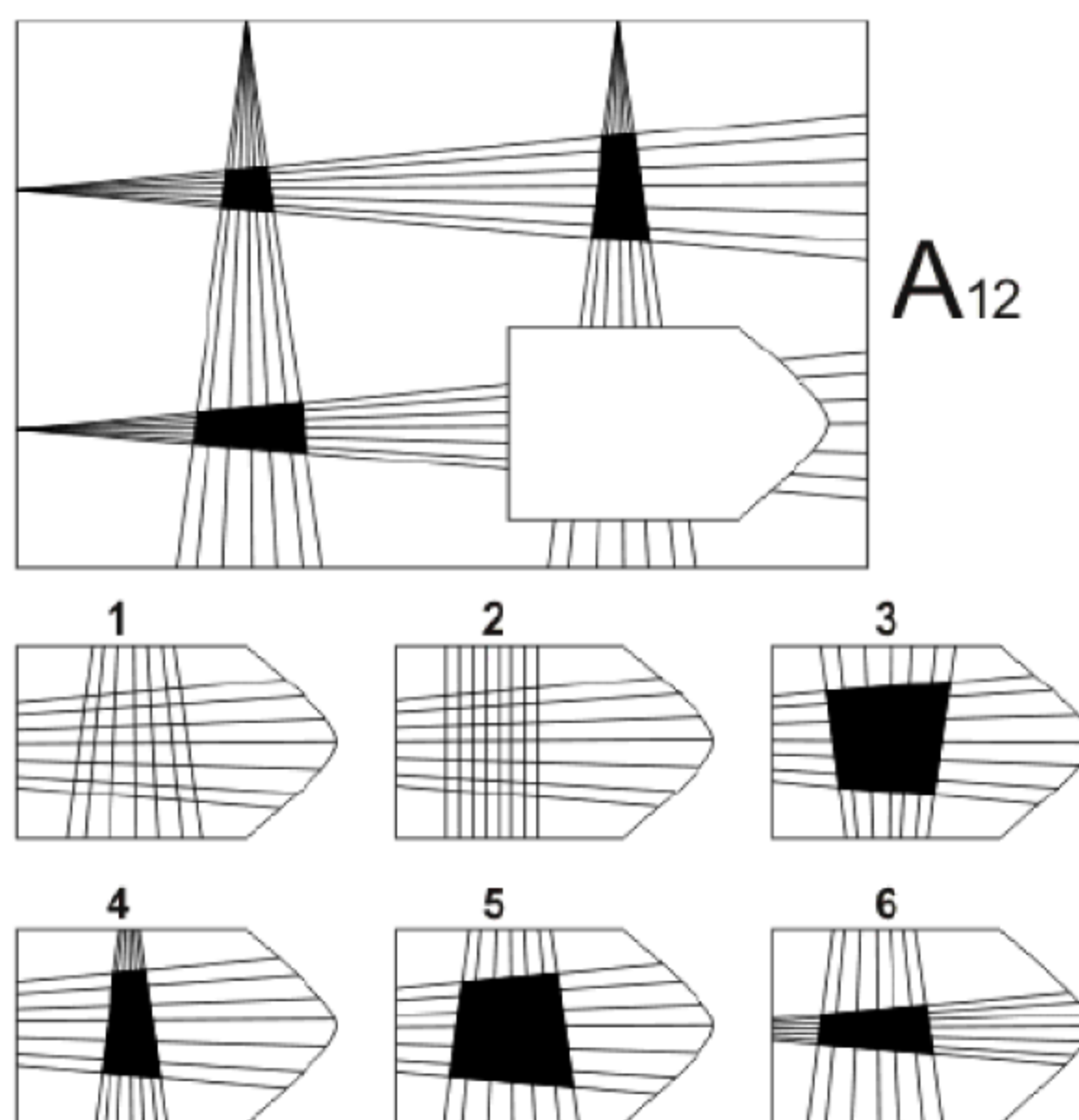
хического заболевания. Процедуру обследования с помощью разработанной им методики Г.Роршах назвал «психодиагностика». До настоящего времени методика Г.Роршаха является широко используемой во всём мире.

Широкое распространение тестов (к концу 1920 г. существовало около 1300 тестов, с помощью которых в течение года измеряли 30 млн показателей) привело к проникновению тестов во все сферы человеческой жизни. Основные направления использования тестов – образование (педология), промышленность и армия (психотехника). Столь интенсивное развитие прикладной психологии привело к её отрыву от академической психологии, имевшему негативные последствия. Результаты тестов были фрагментарными, зачастую противоречивыми, прогнозы по этим результатам не сбывались, для грамотной интерпретации требовалось фундаментальное обоснование методик. Для разрешения этих проблем требовалось, во-первых, развитие академической психологии для определения психологической природы изучаемых явлений и, во-вторых, определение места и значения тестов в изучении этих явлений. Иными словами, каждая применяемая психодиагностическая методика должна иметь под собой научную основу.

В 20-е годы XX века произошли прогрессивные изменения в изучении природы интеллектуальных способностей. В это время были опубликованы работы Чарльза Спирмена, который выявил наличие корреляционных связей между результатами тестирования различных способностей (успешность которых определяется специфическими для каждого вида деятельности S-факторами) и предложил термин «G-фактор», который латентен, но предопределяет успех человека в различных видах деятельности. На основе теории Спирмена был разработан тест Равена (рис. 2.3).

Многие психологи не принимали концепции Спирмена, считая, что факторов, определяющих интеллект, больше. Так, Дэвид Векслер разработал иерархическую модель, состоящую из трёх уровней: уровень общего интеллекта, уровень групповых факторов (вербальные и невербальные интеллектуальные способности) и уровень

специфических факторов. Всемирно известный тест Векслера разработан в соответствии с этой моделью, отдельные субтесты, соответствующие специфическим факторам, сгруппированы в вербальный и невербальный блоки. Уровень общего интеллекта по тесту Векслера определяется как интегративная оценка вербального и невербального факторов. Существуют и другие модели интеллекта, в каждой из которых – своя процедура измерения. Аналогично существуют различные подходы и при диагностике личности. Умение обобщать результаты применения психодиагностических методик с опорой на соответствующие методологические основы, давать комплексную оценку личности, при которой с помощью психодиагностики получена лишь часть информации, является показателем квалификации психолога-психодиагноста.



**Рис. 2.3. Образец задания теста Равена**

В нашей стране психодиагностика начала развиваться на рубеже XIX–XX в.в. под влиянием работ И.М.Сеченова (рефлекторная теория психической деятельности), И.П.Павлова (теория условных

рефлексов), В.М.Бехтерева (объективные проявления психических процессов).

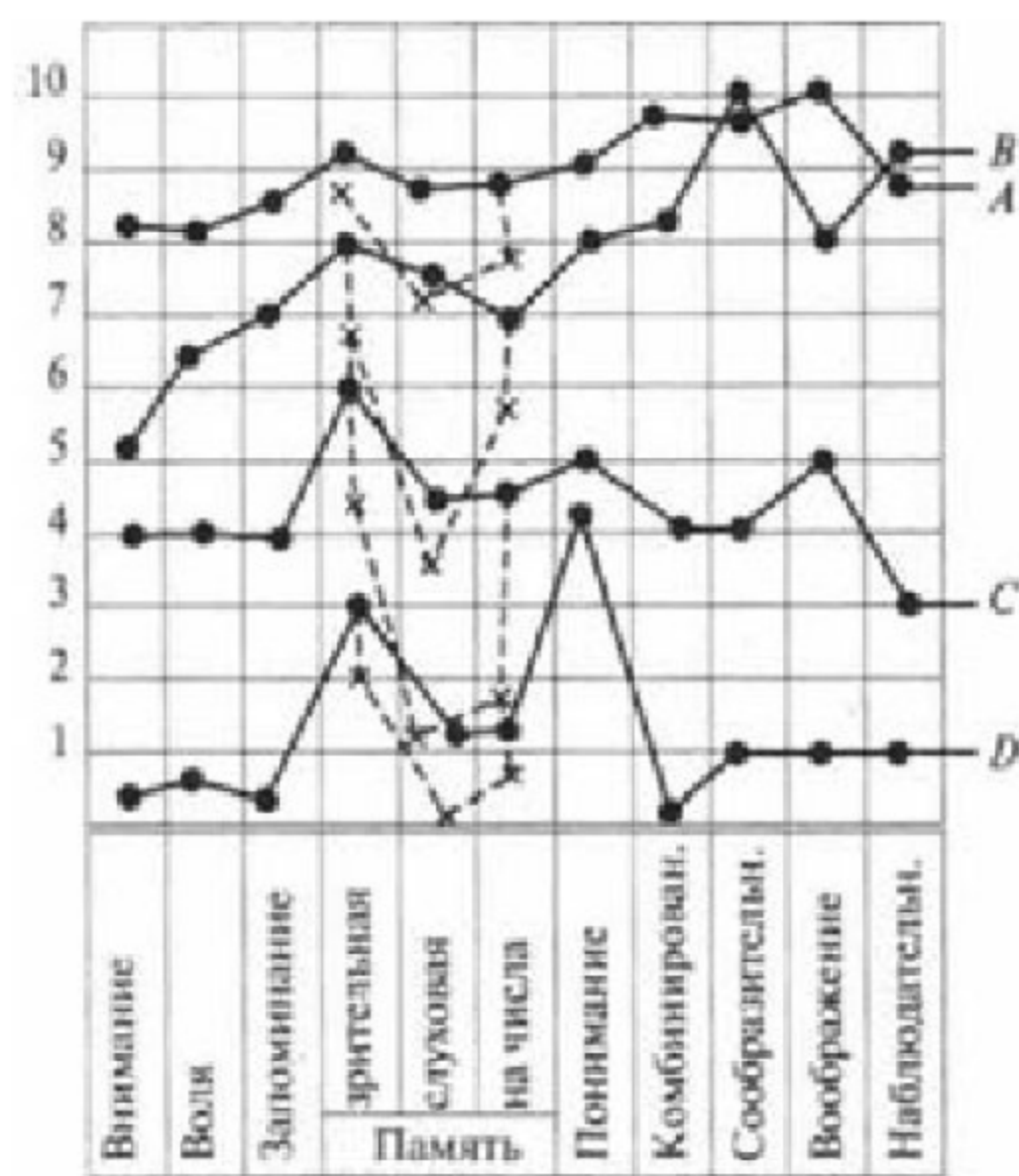
В России начала XX века тесты были хорошо известны и популярны. Особый интерес к тестам проявили педагоги, которые надеялись с помощью тестов оценить степень одарённости учащихся, а также спрогнозировать особенности развития детей. Отметим, что как и за рубежом, в России были как сторонники, так и ярые противники естественнонаучного экспериментального направления в психологии, которые постоянно дискутировали между собой.

Работы по психологическому тестированию в дореволюционной России связаны, в первую очередь, с именами выдающихся учёных Г.И.Россолимо и Ф.Е.Рыбакова.

Одна из первых значительных дореволюционных отечественных работ по психологическому тестированию была выполнена в 1909 г. крупнейшим невропатологом и психиатром **Григорием Ивановичем Россолимо** в Московском университете.

Г. И. Россолимо, имеющий греческие корни, родился 5 декабря 1860 г. в Одессе. В 1884 г. закончил медицинский факультет Московского университета, был однокурсником и близким другом А. П. Чехова. С 1890 являлся заведующим клиникой нервных болезней при клинике внутренних болезней А. А. Остроумова и одновременно занимался преподавательской деятельностью в Московском университете, активно выступает на педагогических и психологических съездах, освещая проблемы изучения обучения и воспитания умственно отсталых детей. Г. И. Россолимо ведёт большую общественную деятельность. Он является одним из организаторов и первым председателем Общества невропатологов и психиатров, основателем и редактором журнала «Невропатология и психиатрия», одним из организаторов психологического общества при Московском университете. В 1911 г. Г.И.Россолимо организовал на свои деньги Институт детской психологии и неврологии, который передал в 1917 г. Московскому университету и продолжал свою работу в качестве профессора МГУ и директора Неврологического института им. А. Я. Кожевникова.

Основной целью работы Россолимо (как и А.Бине) была выработка критериев отличия нормальных детей от имеющих разные степени умственной отсталости. Но, в отличие от А.Бине, Г.И.Россолимо при создании теста шёл не эмпирическим путём, а опирался на свои теоретические представления о структуре личности и интеллекта. Полученная в результате система обследования, названная «методикой индивидуального психологического профиля» сводилась к определению 11 психических процессов, объединённых в три группы: внимание и воля, точность и прочность восприятия, ассоциативная деятельность (рис. 2.4). Задания варьировались по категориям обследуемых (дети, взрослые интеллигентные, взрослые неинтеллигентные).



**Рис. 2.4 Психологические профили по методике Г.И.Россолимо**

Россолимо предложил графическую форму представления результатов тестирования – «психологический профиль», который наглядно демонстрировал соотношение психических процессов испытуемого.

При построении интеллектуального профиля по каждой из 11 групп заданий выставляются оценки (от 0 до 10). Г.И.Россолимо вывел формулу суммарного показателя интеллектуального развития:

$$P=t+m+r\%+a \quad (2.2)$$

где  $t$  – «психический тонус» (внимание, воля);  $m$  – точность и прочность восприятия;  $r\%$  – процент сохранения в памяти;  $a$  – оценка по высшим ассоциативным процессам.

Профили Россолимо получили высокую оценку научной общественности и были переведены на несколько языков.

Для представления на одном графике результатов нескольких различных тестов Г. И. Россолимо ввёл и впервые применил понятие профильной оценки, как способа представления количественных результатов, при котором оценки по отдельным заданиям, составляющим тестовую батарею, приводятся к соизмеримым единицам. В настоящее время применяются следующие приёмы приведения к единой шкале оценок результатов различных тестов: метод выравнивающих коэффициентов, унификация первичных оценок, преобразование стандартизированных оценок в шкалу с едиными значениями среднего значения и стандартного отклонения [11]. Так, например, для результатов интеллектуальных тестов принята шкала со средним значением 100 и стандартным отклонением 16 (рис. 2.1)

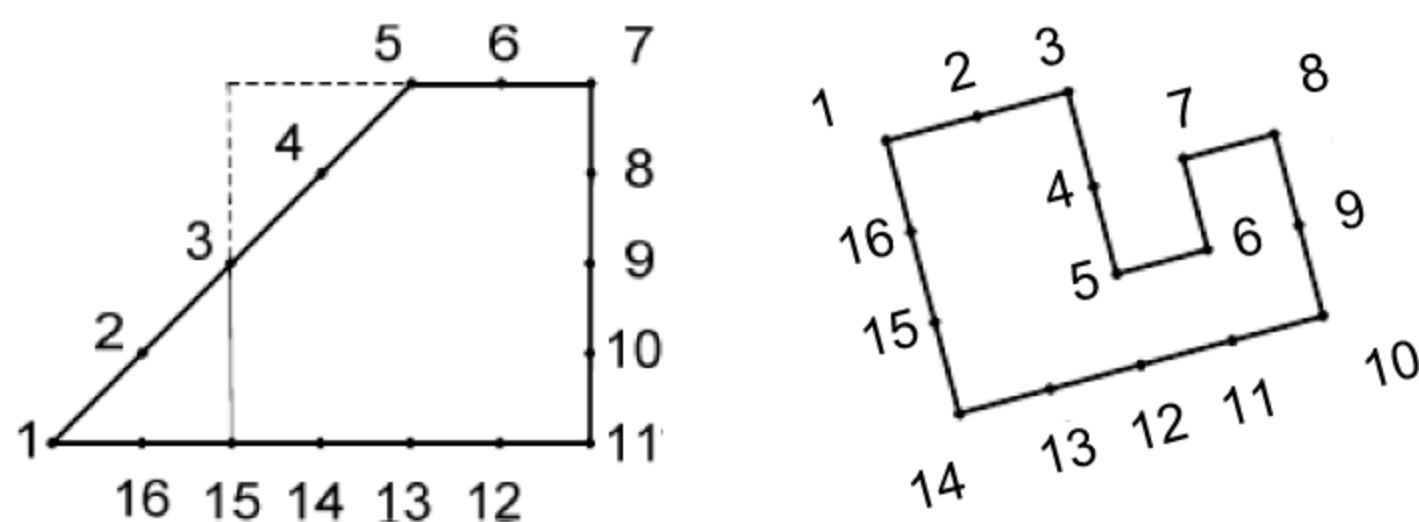


Рис. 2.5 Образец заданий теста Ф.Е.Рыбакова



Не менее интересен и подход Ф.Е.Рыбакова к определению уровня развития пространственного воображения. Его тест «Фигуры Рыбакова», опубликованный в 1910 г., (рис.2.5) получивший широкую известность в мире, в настоящее время входит в качестве одного из субтестов в различные тесты интеллекта. Суть заданий теста сводится к тому, что нужно мысленно разрезать предлагаемую фигуру так, чтобы из полученных двух частей можно было сложить квадрат.

После Октябрьской революции, в 1920-30 годы в нашей стране широко использовались различные тесты, развивались психотехника и педология (прикладная психология в образовании). В 20-е годы в нашей стране значительное развитие получили психология труда и психотехника. В рамках этих отраслей психологии развивалась отечественная психодиагностика, результаты которой нашли применение в промышленности, на транспорте, в профессиональном обучении.

Получила широкое развитие и возрастная психодиагностика, являющаяся составной частью педологии, которая была задумана, как комплексная наука, занимающаяся целостным изучением детей. Многие известные психологи работали в области психологического тестирования (М.Я.Басов, П.П.Блонский, Л.С.Выготский). Наряду с прикладными исследованиями, развиваются и теоретические представления о психодиагностике. Так, Л.С.Выготским разработано учение о психологическом диагнозе. Вместе с тем, общий кризис психодиагностики, связанный с чрезмерным увлечением тестированием и отрывом от академической науки, имел место и в нашей стране. Итак, во-первых, теория психологического тестирования существенно отставала от практики, и, во-вторых, в связи с широким распространением тестов методики часто попадали в руки людей, не имеющих должной психологической подготовки. Таким образом, издержки практики тестирования отождествлялись с недостатками тестирования в целом.

Усугубило положение и то, что результаты психодиагностических обследований (особенно детей) шли вразрез с марксистско-



ленинской идеологией. Так, М.Ю.Сыркиным был доказан социальный аспект существования интеллектуальных различий между слоями и классами общества. Также обследования конца 20-х годов показали, что IQ советских учащихся в среднем на 7% ниже, чем американских школьников.

Кроме социальных различий в интеллекте детей, были доказаны и национальные различия. Широкую известность приобрели исследования интеллекта детей в Узбекистане, показавшие, что уровень интеллекта узбекских детей значительно ниже, чем русских. Причина – тяжёлые социально-бытовые условия и особенности психофизиологического развития (как известно, в среднеазиатских семьях традиционно много детей).

Также было доказано, что интеллектуальные различия существуют и у славянских детей: у детей служащих интеллект выше, чем у детей рабочих, а самый низкий уровень – у детей крестьян. Эти различия носили устойчивый характер. Исследования взрослых показали сходные результаты, причём самый высокий интеллект оказался у выходцев из дворянства, духовенства и купечества. Всё это послужило причиной критики тестов. Некоторые идеологи выступали с высказыванием, что «показатель интеллекта в Советской России не ценится так высоко, как на Западе, у нас существуют другие критерии способности членов общества к той или иной деятельности». Естественно, здесь речь шла о преданности партии, лояльности к власти и т.д. Как отмечал американский психолог О. Брим, залогом позитивного отношения к тестам является одобрение обществом принципа конкуренции между его членами, в соответствии с которым ведущие роли должны быть отданы наиболее талантливым людям. В Советской России 30-х годов об этом не могло быть и речи.

В начале 30-х годов началась принципиальная критика педологии, завершившаяся в 1936 году Постановлением «О педологических извращениях в системе наркомпросов», которым было запрещено применение тестов в школе. По существу, были прекращены

все психодиагностические исследования. Только через 40 лет это направление исследований было восстановлено в своих правах.

В 1960-70-е годы в советской психологии развивается «качественный подход» (в противовес количественному подходу в западной тестологии). В рамках качественного подхода та или иная способность индивида изучалась в лабораторных условиях при выполнении соответствующего вида деятельности. Этот метод, хотя и был достаточно эффективным для решения исследовательских экспериментально-психологических задач, но в силу своей громоздкости и ориентированности на лабораторные условия не мог заменить тесты.

Использование тестов отечественными психологами возобновилось только в 70-е годы в медико-психологических исследованиях, исследованиях по возрастной психологии, в профориентации и профотборе, а также в судебно-психологической экспертизе. В этих исследованиях, осуществляемых без поддержки, а часто и при критике со стороны официальной психологии, во-первых, накапливался опыт организации и проведения диагностических исследований, а во-вторых, были получены данные, дополняющие известные теоретические представления об индивидуально-психологических особенностях личности в норме и патологии. К концу 70-х годов был опубликован ряд монографий, посвящённых как психодиагностике в целом, так и отдельным её проблемам. Психодиагностика тем самым была признана одной из основных сфер приложения профессиональных возможностей психологов и курс психодиагностики стал одним из базовых курсов подготовки специалистов-психологов.

Результатом того, что в 30-е годы были ликвидированы все психодиагностические разработки, связывающие науку и практику в психологии, стало фактическое отсутствие отечественных психодиагностических методик. Начиная с 80-х годов в отечественной психологической практике начинается период активного применения зарубежных методик, в большинстве своём не адаптированных и не стандартизированных, представляющих собой непрофессио-

нально (с психологической точки зрения) выполненные переводы. У таких методик очень часто отсутствует методологическая база, научный анализ теоретических конструктов заменяется простым описанием процедуры применения методики. Например, понятие «интеллект» применительно к психодиагностической методике должно быть расшифровано: что понимает автор под понятием «интеллект», какие стороны интеллекта измеряются с помощью методики, на основе каких научных концепций разработана система заданий и т.п. Кроме того, должны быть данные о надёжности и валидности методики, а также сведения о выборке стандартизации и нормы.

Вместе с тем, в 80-е годы прошлого века в отечественной психодиагностике начинают появляться варианты адаптированных и стандартизированных зарубежных методик. Широкую известность получили адаптированный вариант теста Д.Векслера под ред. А.Панасюка, различные варианты адаптации опросника MMPI (*Minnesota Multiphasic Personality Inventory*) под ред. Л. Н. Собчик и под ред. Ф.Березина, адаптированные варианты психодиагностических методик Р. Кеттелла, М. Люшера, Т. Лири. Одновременно разрабатывались и отечественные психодиагностические методики: патохарактерологический диагностический опросник А. Е. Личко (ПДО), Школьный тест умственного развития, созданный под руководством К. М. Гуревича (ШТУР), диагностические методики для детей, выполненные под руководством Л. А. Венгера и О. М. Дьяченко.

Одна из существенных проблем отечественной психологической диагностики конца XX века – острая нехватка психодиагностического инструментария для практических психологов, тестовый материал зачастую изготавливался кустарно, с помощью многократного ксерокопирования. В настоящее время эта проблема частично решена с созданием в Петербурге предприятия ИМАТОН, которое занимается подготовкой и продажей психодиагностических методик для психологов, а также организует научное сопровождение

этих методик, ведёт работы по стандартизации, проводит повышение квалификации пользователей этих методик.

Итак, история отечественной психодиагностики развивалась от массового чрезмерного использования тестов на этапе становления, через длительный период запрета на их применение, до второго рождения во второй половине XX века, периода широкого и неконтролируемого использования западных тестов и постепенного развития научной базы психодиагностики в настоящее время.

### **Контрольные вопросы и задания**

- 1. Какой вклад внесли исследования Ф.Гальтона в развитие методов психологической оценки?*
- 2. Проведите сравнительный анализ тестов А. Бине и Д. Векслера.*
- 3. Охарактеризуйте развитие психологической диагностики в России на протяжении XIX-XXI в.в.*

## **2.2 История развития системы оценивания в образовании**

Система оценивания в образовании имеет богатую историю и свои традиции. Впервые оценивать знания в России начали в 1737 году, при императрице Анне Иоановне. Оценки были словесными, семантически многообразными в пределах полярных понятий «хорошо – плохо». Позднее появился количественный эквивалент оценки – отметка. В 1837 году была официально утверждена 5-балльная система оценки: «1» – слабые успехи; «2» – посредственные; «3» – достаточные; «4» – хорошие; «5» – отличные. К 1917 г. эта система официально существовала во всех типах учебных заведений России: в гимназиях, реальных училищах, духовных семинариях, земских и церковно-приходских школах и т.д. После Октябрьской революции вместо прежних типов учебных заведений была создана единая трудовая школа, в которой труд выступает и как учебный предмет, и как метод учебно-воспитательной работы. В мае 1918 г было принято решение Наркомпроса РСФСР «Об отмене отметок». Были отменены и экзамены. Широко практиковались выставки работ учащихся, отражавшие результаты их учебной

деятельности. На выставках были и натуральные предметы, и модели, всевозможные таблицы, диаграммы, чертежи, рисунки, которые давали наглядное представление о том, чего достигли ученики, выполнив задание, закончив тему. Перевод из младшей группы в старшую проводили на основании отзыва педагогического совета школы об успехах учащихся. Практиковалась также коллективная аттестация учебной группой своих товарищей. Обсуждение было взыскательное, принципиальное; при плохой аттестации кого-либо из товарищей требовали даже оставления его на повторный год обучения в той же группе.

Такой «коллективный» подход к оценке знаний не позволял оценить результаты учебной деятельности каждого ученика, у многих учащихся пропадал стимул к учебной деятельности, поскольку оценивались средние показатели по классу. Многие учителя неформально вернулись к прежней, дореволюционной системе оценки. Реальное положение дел закрепилось в 1935 году постановлением СНК СССР и ЦК ВКП (б) в котором была восстановлена пятибалльная дифференцированная словесную систему оценок и отметок знаний, умений и навыков учащихся: «отлично», «хорошо», «удовлетворительно», «плохо» и «очень плохо». Из-за смысловой неопределенности оценка «удовлетворительно» несколько позже была переименована на «посредственно».

С 1944 г. во всех школах страны перешли на количественную пятибалльную систему отметок: 5 – отлично, 4 – хорошо, 3 – посредственно, 2 – плохо, 1 – очень плохо. Были разработаны и критерии оценок и отметок, что стандартизировало систему оценивания в различных типах учебных заведений и позволяло учителю более точно определять уровень знаний и умений учащихся.

Пятибалльная система оценивания не является мировой общепринятой системой. В большинстве стран мира действует более широкая шкала оценок - от 6 (Польша) до 100-балльной (Япония). В Белоруссии и Латвии используется 10-балльная система, на Украине – 12-балльная. Во Франции действует 20-балльная. В Германии и Чехии принята 5-балльная система, но оценки ставятся в об-

ратном порядке и наивысшей оценкой является единица. В США также принята 5-балльная система оценки, но вместо цифр там используют буквы - A, B, C, D, F. В Великобритании принята словесная оценка работы ученика.

Метод оценивания знаний учеников в различных странах имеет различную степень субъективности и зависимости от личности и установок учителя. Повышение объективности итоговой оценки связано с внедрением в образовательный процесс образовательных тестов. Наибольшее распространение эти тесты получили в США (в 1845 году в Бостонской школе перевели экзамены с устной формы на письменную, что в какой-то мере стандартизировало процедуру экзамена), где тестированию подвергаются все дети с 5 лет. В XX веке тесты достижений, по которым оценивается уровень обученности школьников различных классов, стали применяться повсеместно. В нашей стране, начиная с 2004 года, введена обязательная аттестация выпускников общеобразовательных школ в форме единого государственного экзамена (ЕГЭ).

В настоящее время образовательной мобильности большое значение имеет сопоставление качества образования в различных странах. Мониторинговые исследования качества образования в различных странах проводят Международная ассоциация по оценке учебных достижений IEA (International Association of Evaluation of Educational Achievements) и Организация экономического сотрудничества и развития - OECD (Organisation for Economic Cooperation and Development). В списке IEA - исследования по изучению качества и понимания текста - PIRLS, по оценке качества математического и естественно-научного образования - TIMSS, по информационным технологиям в обучении - SITES и по качеству гражданского образования – ICCS. Основная задача этих исследований – оценить качество образования в этих областях в соответствии с теми программами и задачами, которые разные страны мира ставят перед своими системами образования. В TIMSS, например, изучается подготовка учащихся на ключевых этапах обучения: по



окончании начальной школы и почти в конце основной школы, в восьмом классе.

В исследованиях PISA выделено 6 уровней подготовки учеников. Особое место занимает 2-й уровень, являющийся базовым. Например, применительно к показателю качества понимания текста (PIRLS) считается, что ученик достиг базового уровня, если он способен найти в тексте одну или несколько единиц информации, требующих дополнительного, но не сложного осмысления, распознать главную мысль, понять связи отдельных его частей и дать интерпретацию этим частям. На 6-м уровне ученик способен детально и точно интерпретировать текст в целом, каждую единицу информации, сообщенной в самых глубинных слоях текста. То есть он понимает, что хотел сказать автор, умеет сопоставлять разные точки зрения, высказывать свою, аргументировать её [31].

Все перечисленные подходы к педагогической оценке относятся к «знаниевой парадигме», т.е. к контролю и оценке, в первую очередь, знаний обучающихся. Как отмечает А.М.Новиков [57, с.414], сложились даже стандартные выражения, над смыслом которых уже мало кто задумывается: «школа знаний», «поход за знаниями», «общество знаний» и т.д. Сегодня, в социально-экономических и информационно-технологических условиях новой эпохи и соответствующему ей проектно-технологическому типу организационной культуры оцениваться должны не только знания, но и умения, компетенции. Результат ЕГЭ, на основании которого абитуриенты зачисляются в вуз, по большому счёту, никак не характеризуют потенциальные возможности будущего специалиста. Образовательные стандарты общего и профессионального образования сформулированы в терминах компетенций. Всё это свидетельствует о необходимости разработки принципиально иных подходов к оценке результатов образования.

В 1990-е гг. в различных учебных заведениях стали апробироваться новые подходы к оцениванию познавательной деятельности школьников. Они заключались в переходе с пятибалльной на многобалльную систему (от 100 до 1000), известную как модульно-

рейтинговая. Смысл ее заключался в том, что каждое задание в зависимости от сложности, нестандартности, креативности оценивалось по определенной, пропорционально увеличивающейся шкале. При этом ученик мог выбрать задание в соответствии с самооценкой, уровнем притязаний и амбиций. Как показал опыт применения такой системы оценивания, она, особенно на первых порах, вызывает интерес у школьников, стимулирует их активность. Положительной стороной является и включение механизма самооценки учащихся, их самоопределения. В некоторых образовательных учреждениях для повышения мотивации вводились «школьные деньги», которые ученики получали в зависимости от количества набранных баллов. На заработанные «деньги» ученики могли приобрести в школьном магазине вполне реальные и престижные для подростков вещи. Однако все же модульно-рейтинговая система и ее многочисленные модификации представляли собой систему, конвертируемую в итоге в стандартный вариант отметок. В результате набранные учениками сотни и тысячи баллов переводились в традиционную пятибалльную шкалу.

В течение последних десятилетий разрабатывались и апробировались также различные методики оценивания деятельности школьников по конечным результатам. Одной из таких форм является сдача старшеклассниками в конце полугодия всего освоенного материала в ходе зачетной недели и соответственно получение «зачета» или «незачета». Преимущество данного подхода в его аналогичности вузовской системе обучения, на которую предстоит перейти большинству старшеклассников. Однако на практике наблюдались и характерные для такой системы оценивания недостатки – ученики по-студенчески «откладывали все на потом» и тщетно пытались в течение зачетной недели овладеть непосильным для них объемом материала.

Все вышеперечисленные методы педагогической оценки суть методы контроля и оценки результатов обучения со стороны педагогов. Вместе с тем, **в условиях непрерывного образования, «образования в течение всей жизни» одним из важнейших качеств**

**становится самоконтроль и самооценка своей учебной деятельности.** Поэтому обучение методам самоконтроля и самооценки обучающимися своей учебной деятельности является сегодня актуальным направлением развития методов оценки в образовании.

В конце 90-х гг. в некоторых образовательных учреждениях нашей страны стали применяться субъект-субъектные подходы к оцениванию знаний, получившие к тому времени широкое распространение в США. Эти подходы включают непрерывную оценку и самооценку посредством интеграции качественной и количественной ее составляющих, перевода внешней субъективной оценки во внутреннюю объективную самооценку и формирование способности обучаемого к непрерывной самооценке в процессе непрерывного самообразования.

Такая система получила название учебное «портфолио». Суть ее заключается в том, что организуется накопление, отбор и анализ образцов и продуктов учебно-познавательной деятельности ученика, а также соответствующих информационных материалов из внешних источников (от учителей, одноклассников, родителей). На основе этого проводится всесторонняя количественная и качественная оценка уровня обученности данного ученика и дальнейшая коррекция процесса обучения. При этом делается акцент на то, что ученик знает, умеет, а не на то, чего он не знает и чего не умеет.

В настоящее время в условиях введения новых образовательных стандартов наметилась тенденция объединения педагогической и психологической оценки. Результатом образования является система знаний, умений, навыков и освоенных компетенций. Оценить компетенции обучающихся можно только путём сочетания строго-формализованного и малоформализованного методов психологической оценки.

### **Контрольные вопросы и задания**

- 1. Перечислите достоинства и недостатки использования тестов педагогической оценки в форме тестирования*
- 2. Какая система оценивания результатов образования представляется Вам наиболее эффективной?*

## Глава 3.

### Строгоформализованный метод психологической оценки

#### 3.1. Классификация строгоформализованных методик

Методы психологической оценки, как уже отмечалось в параграфе 1.3, подразделяются на две группы (рис. 1.3): малоформализованные методы и методы высокого уровня формализации, которые конкретизируются в строгоформализованных методиках: тестах; опросниках; психосемантических методиках, методиках проективной техники; психофизиологических методиках. Для этих методик характерны: **стандартизация процедуры проведения** (регламентация); **стандартизация оценки выполнения, надёжность и валидность**.

Следует отметить, что по сложившейся традиции многие авторы называют словом «тест» методики, относящиеся к другим классам строгоформализованных методик. Например, широко известный 16-факторный личностный опросник Р.Кеттелла называют «тест Кеттелла», проективную методику, разработанную Г.Роршахом, «тест Роршаха», существуют также «тест Люшера», «тест Розенцвейга» и т.д. Это устоявшиеся названия, и при их использовании необходимо помнить о данном терминологическом нюансе.

Любая строгоформализованная методика (СФМ) представляет собой объективное и стандартизованное измерение образцов или проб поведения. Психологические тесты проводятся на малой выборке тщательно отобранных образцов поведения индивидуума. Анна Анастази проводит здесь аналогию с биохимиком, который делает заключения о составе крови пациента или качестве питьевой воды на основании одной или нескольких взятых проб [4].

Соответствие СФМ изучаемому аспекту поведения зависит от содержания и количества заданий. Диагностическая или предска-

тельная ценность СФМ зависит от того, насколько она может служить индикатором измеряемой области поведения. Отметим, что задания, из которых сконструирована психодиагностическая методика (в отличие от тестов достижений, применяющихся в качестве методов оценки в образовании), не обязательно должны иметь близкое сходство с прогнозируемым поведением. Так, в невербальных интеллектуальных методиках (тест Равена), либо в проективных методиках (тест Роршаха) степень соответствия характеристик деятельности человека в ситуации тестирования и в других ситуациях (как в настоящем, так и в будущем) определяется различными типами валидности, которые будут рассмотрены в параграфе 3.2.

Как уже отмечалось в главе 1, психодиагностический метод конкретизируется в трёх основных диагностических подходах, которые исчерпывают всё множество строгоформализованных психодиагностических методик. В рамках объективного подхода разрабатываются тесты и психофизиологические методики, в рамках субъективного подхода – опросники и психосемантические методики, в рамках проективного подхода – методики проективной техники. Далее будут рассмотрены подходы к классификации перечисленных методик.

Существует несколько классификаций строгоформализованных методик по форме.

**А. Классификация по форме проведения испытания:** индивидуальные и групповые методики.

В ходе индивидуального обследования взаимодействие экспериментатора и испытуемого происходит один на один. Начиная с армейских тестов «Альфа» и «Бета» разработано достаточно много тестов, опросников и психосемантических методик, позволяющих обследовать одновременно группу людей.

Следует отметить, что в ходе группового испытания гораздо сложнее отследить состояния испытуемого, которые могут повлиять на результат (болезнь, утомление, беспокойство, тревожность). Как правило, лица не знакомые с процедурой группового тестирования, оказываются в менее выгодном положении. Поэтому в тех

случаях, когда принимаемое по результатам тестирования важно для испытуемого, желательно дополнять результаты группового обследования тестирования индивидуальным.

Приведём пример. Высокий результат, полученный испытуемым по результатам группового обследования интеллектуальных способностей с помощью теста Равена, свидетельствует о хорошем уровне развития целостного восприятия, произвольного внимания, оперативной памяти, различных видов мышления. Если же испытуемый получает низкий балл, должна быть дополнительная, желательно индивидуальная проверка интеллектуальной сферы с помощью других методик.

Обследование с применением методик проективной техники в подавляющем большинстве случаев желательно проводить только индивидуально, так как при групповом обследовании сложнее наблюдать за поведением испытуемого при выполнении задания, фиксировать его высказывания, паузы и другую информацию, играющую определяющее значение при составлении диагноза. Отдельные категории лиц (дети младшего возраста, пожилые люди, испытуемые с ограниченными интеллектуальными или физическими возможностями), обследование которых можно проводить только индивидуально.

Существуют определённые правила проведения группового обследования:

1. Количество испытуемых в группе должно определяться возрастными нормами группового обследования: дети дошкольного возраста – 4-6 человек, школьники – до 15 человек, взрослые – до 30-50 человек.

2. При проведении группового обследования у экспериментатора должно быть достаточное число ассистентов (не менее 1-2 ассистента на группу дошкольников или школьников, не менее 1 ассистента на 10 взрослых испытуемых).

3. Во время проведения обследования испытуемые должны сидеть по одному за столом (если обследование проходит в большой



аудитории, испытуемые должны сидеть на расстоянии не менее 1 м друг от друга).

4. Каждый испытуемый должен быть обеспечен индивидуальным набором методик. При проведении обследования с помощью опросников каждый испытуемый должен быть обеспечен буклетом с текстом опросника (желательно выполненным типографским способом).

5. При групповом обследовании необходимо строго придерживаться инструкции, соблюдать единообразие условий проведения эксперимента.

**Б. Классификация по форме ответа:** устные и письменные методики.

Устными чаще всего бывают методики, разработанные исключительно для индивидуального обследования. Методики, при выполнении которых испытуемому требуется самостоятельно заполнять бланки, называются письменными и могут применяться как при индивидуальном, так и при групповом обследовании. Как устные, так и письменные ответы могут либо формулироваться испытуемым самостоятельно (открытые ответы), либо выбираться из предложенных ответов (закрытые ответы). В письменных методиках ответы даются испытуемым в тестовой тетради или на отдельном бланке ответов. Для анализа устных ответов испытуемого целесообразно использовать диктофон.

**В. Классификация по типу инструментария,** который используется при тестировании: бланковые, предметные, аппаратные, компьютерные.

– Бланковые (тесты «карандаш и бумага»): бланк или тетрадь, в котором содержится инструкция по применению, примеры решения заданий, сами задания и графы для ответов (для ответов может быть отдельный бланк, что позволяет использовать одни и те же тестовые тетради многократно).

– Предметные: диагностический инструментарий представлен в виде предметов: кубиков, карточек. Обследования с использованием предметных методик чаще всего проводятся индивидуально, к их числу относятся многие проективные методики (тесты Роршаха, Люшера, Сонди, ТАТ и др.), а также интеллектуальные тесты (субтесты теста Кеэса, кубики Коса и др.).

– Аппаратурные тесты – это тип методик, требующий применения специальных технических средств или оборудования для проведения исследования или регистрации полученных данных. Это приборы для исследования времени реакции, изучения особенностей восприятия, памяти, мышления. В последнее практически все аппаратурные тесты компьютеризированы.

– Компьютерные тесты – автоматизированный вид тестирования в форме диалога испытуемого с персональным компьютером. Многие известные методики в настоящее время имеют компьютерную версию (ММРІ, 16-ФЛО, тест Люшера). Эти версии, строго говоря, требуют дополнительных проверок на надёжность и валидность и выработки статистических норм.

В следующих параграфах будет рассмотрена специфика и приведены примеры основных типов СФМ.

### 3.1.1 Тесты

Тест – это стандартизированная методика психологического измерения, предназначенная для диагностики выраженности у индивида психических свойств или состояний при решении практических задач.

Отличительная особенность тестов состоит в том, что **они состоят из заданий, на которые от испытуемого нужно получить правильный ответ**. В опросниках, проективных и психофизиологических методиках правильного ответа нет.

По характеру стимульного материала тесты делятся на вербальные и невербальные. В вербальных тестах основным содержанием работы испытуемых являются операции с понятиями, мыслительные действия, осуществляемые в словесно-логической форме. Эти

тесты очень чувствительны к различиям в языковой культуре, уровню образования, профессиональным особенностям.

Ниже в качестве примера приведены образцы заданий из вербальной шкалы детского варианта теста Д.Векслера. Эти задания стали классическими. Подобные типы заданий в настоящее время включаются во многие интеллектуальные тесты.

### **Субтест 1. «Осведомлённость».**

18. Что такое SOS?
19. Каков примерно средний рост взрослого человека?
20. Где расположена Италия?
21. Сколько килограммов в центнере?

### **Субтест 2. «Понятливость».**

1. Что ты будешь делать, если порежешь себе палец?
6. Почему лучше строить дом из кирпича, чем из дерева?
14. Почему нужно выполнять обещания?

### **Субтест 3. «Арифметический»**

4. Если разрезать яблоко пополам, сколько будет частей?
13. Каким будет число, если  $\frac{2}{3}$  его равняются 36?
15. Если проезд на такси стоит 20 рублей за первую четверть километра и по 5 рублей за каждую последующую четверть километра, то сколько надо заплатить за проезд 2-х километров?

### **Субтест 4. «Сходство» (Аналогии)**

1. Лимоны кислые, а сахар...
4. Нож и кусок стекла оба...
13. Что общего между словами «соль» и «вода»?
16. Что общего между числами «49» и «121»?

### **Субтест 5. «Словарный»**

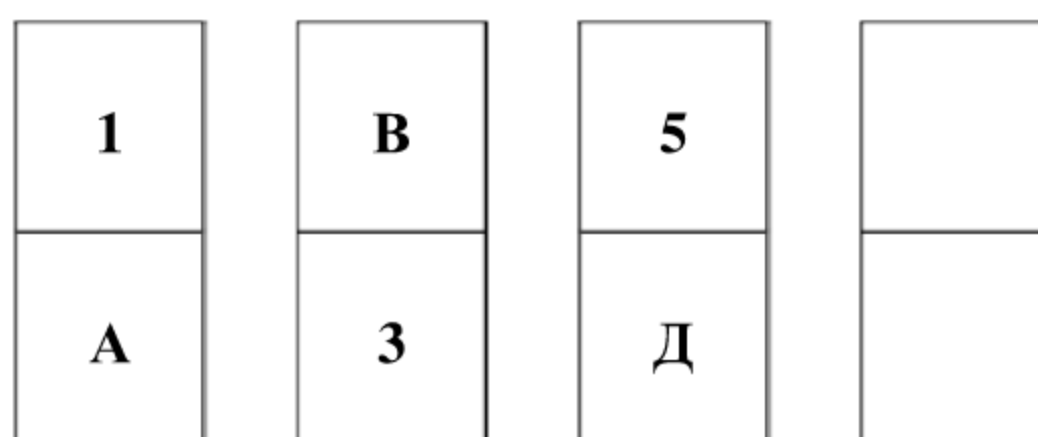
Объясни, пожалуйста, что такое ...

1. Велосипед.
5. Зонтик.
22. Басня.
39. Ирония.

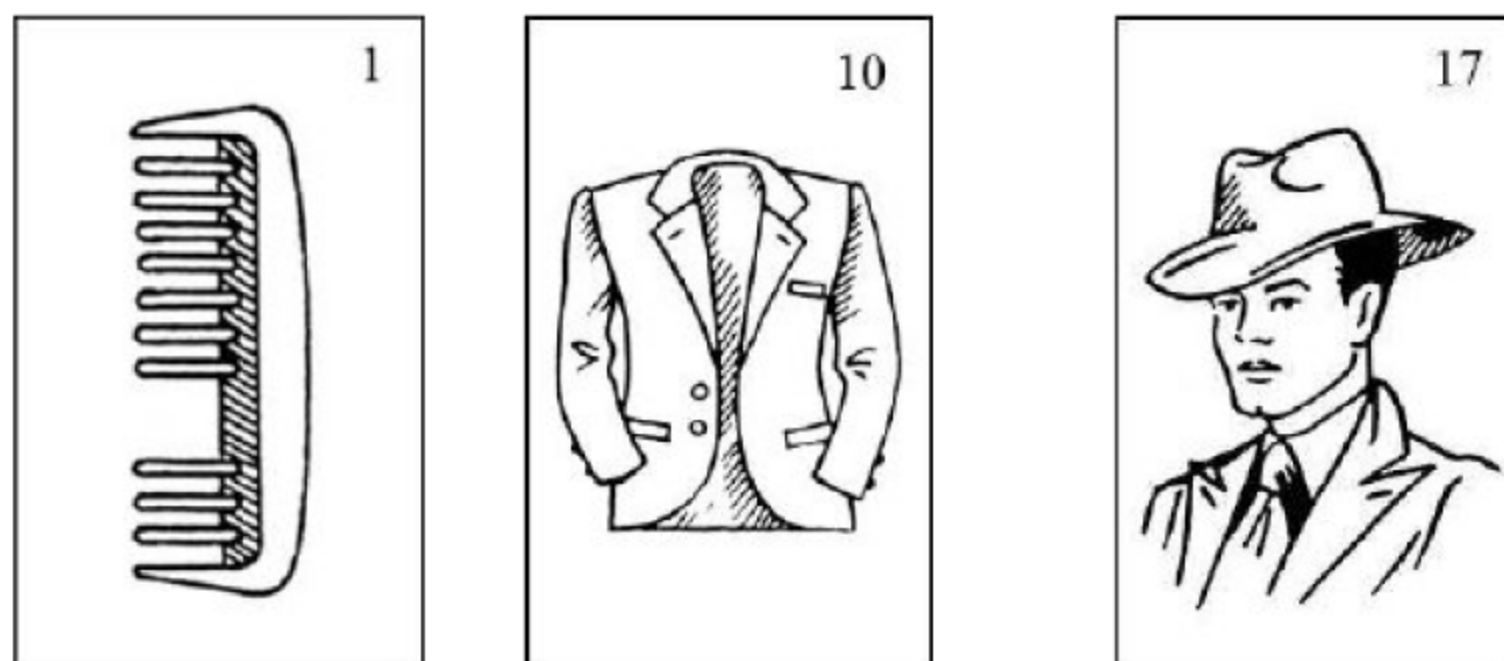
В невербальных тестах тестовый материал представлен в наглядной форме (картинки, чертежи, графические изображения). От

испытуемого требуется понимание вербальных инструкций, само же выполнение заданий опирается на перцептивные и моторные функции. Невербальные тесты уменьшают влияние языковых различий на результат испытания, могут использоваться для тестирования испытуемых с нарушением речи, слуха или с низким уровнем образования. В качестве отдельных субтестов они включены во многие тесты интеллекта, способностей и достижений.

В качестве примеров невербальных интеллектуальных тестов приведём тест Равена (образец задания представлен на рис. 2.3), задания теста Г.Айзенка (рис. 3.1) и задания невербальной шкалы теста Д.Векслера (рис. 3.2 – 3.3).



**Рис. 3.1 Образец одного из заданий IQ теста Г.Ю.Айзенка**



**Рис. 3.2 Образцы заданий субтеста «Недостающие детали» теста Д.Векслера**



**Рис. 3.3 Образцы заданий субтеста «Последовательные картинки» теста Д.Векслера**

По содержанию тесты подразделяются на:

- тесты интеллекта;
- тесты способностей (общих и специальных);
- тесты личности;
- тесты достижений.

### **3.1.2 Опросники**

Опросники – это группа психодиагностических методик, в которой задания представлены в виде вопросов и утверждений. Они предназначены для получения данных со слов обследуемого. В отличие от тестов, в опросниках не может быть правильных и неправильных ответов. Они лишь отражают отношение человека к тем или иным высказываниям, меру его согласия или несогласия.

По степени охвата личностных характеристик опросники подразделяются на:

- одномерные, направленные на измерение какого-либо одного свойства или качества (например, шкала тревожности Тейлора);

– многомерные, направленные на измерение более чем одного свойства или качества (например, 16-факторный личностный опросник Р.Кеттелла).

**По направленности** опросники подразделяются на опросники черт личности, типологические, мотивов, интересов, ценностей, установок.

В соответствии с **принципом, положенным в основу конструирования**, опросники подразделяются на факторные и эмпирические.

В основе построения **факторных опросников** лежит факторно-аналитический принцип. В отличие от эмпирических опросников, в них содержатся задания, которые наиболее тесно связаны с некоторым указанным фактором (факторами). В окончательный вариант опросника входят только те задания, которые непосредственно связаны с общим фактором. Примеры: 16-факторный личностный опросник Р.Кеттелла (16ФЛО или 16PF – Sixteen Personality Factor Questionnaire), личностные опросники Г.Айзенка (EPQ – Eysenck Personality Questionnaire). В качестве примера изложим процедуру разработки опросника 16PF.

В ходе создания опросника Р. Кэттелл опирался на L-данные (life record data), полученные путем регистрации реального поведения человека в реальной жизни. Выделенные Г. Олпортом и Х. Олдбергом 4,5 тыс. слов, обозначающих черты личности и особенности поведения (на базе словаря из 18 тыс. слов), Р. Кэттелл разбил на синонимические группы и отобрал в каждой из них по одному слову, выражающему основное смысловое содержание соответствующей группы. Это позволило сократить список личностных черт до 171. Затем каждая из этих характеристик личности оценивалась экспертами с целью выбора наиболее значимых. Взаимная корреляция экспертных оценок позволила выделить 36 корреляционных плеяд, состоящих из тесно связанных между собой характеристик (при этом связи были как прямыми, так и обратными). Все плеяды содержали пары членов, противоположные по смысловому содержанию, например: веселый – печальный, разговорчивый



– молчаливый и т. д. Так был получен набор из 36 биполярных названий, который был расширен до 46 за счёт специальных терминов, найденных в работах других исследователей. В результате факторизации L-данных было получено от 12 до 15 факторов. В дальнейшем из-за трудностей экспертного оценивания Р. Кеттелл перешел к Q-данным (questionnaire data), т.е. к данным, полученным с помощью опросников. При этом сбор Q-данных координировался с имеющимися L-данными. Р. Кеттеллом созданы разные модификации факторных моделей с различным числом входящих в них факторов, однако, наиболее известным является 16-факторная модель – Шестнадцатифакторный личностный опросник (16ФЛО).

Факторы личности, диагностируемые 16ФЛО, обозначаются буквами латинского алфавита, а для тех факторов, которые выделены на основе Q-данных, используется буква Q: Q1, Q2, Q3, Q4. Факторы имеют бытовые и профессиональные названия. Первые представляют собой общедоступные определения, ориентированные на непрофессионалов. Например, фактор А – «сердечность, доброта – обособленность, отчужденность».

Профессиональные названия предназначены для специалистов-психологов и тесно связаны с научно установленным значением фактора. Тот же фактор А будет определяться как «аффектотимия – сизотимия». Как бытовые, так и профессиональные названия факторов даются в биполярной форме, чем устраняется двусмысленность в определении их содержания.

К недостаткам теоретического подхода Р. Кеттелла относят выраженный эмпиризм, пренебрежение какими-либо исходными теоретическими положениями о содержании и количестве определяемых черт личности [11]. Функциональные связи между переменными выражаются лишь в виде корреляций – меры линейной взаимосвязи между переменными. Не удалось верифицировать и исходную гипотезу об идентичности структурных элементов в факторах, выделенных на основе L-данных и тех, которые были выделены на основе Q-данных. Это приводит к выводу, что Р. Кэттелл

принял факторный анализ за гораздо более эффективный инструмент познания личности, чем он фактически является.

Опросник 16PF (16 Personality Factors Questionnaire) был впервые опубликован Р. Кэттеллом в 1950 г. Были разработаны две основные параллельные формы опросника – А и В, состоящие из 187 вопросов, а также формы С и D – 107 вопросов. Впоследствии были разработаны формы опросника CPQ (Children Personal Questionnaire) для детей от 8 до 11 лет и HSPQ (High School Personal Questionnaire) для подростков от 12 до 17 лет. Приведём примеры вопросов методики 16PF:

1. Я думаю, что моя память сейчас лучше, чем раньше:  
а) да б) трудно сказать с) нет
2. Я мог бы счастливо жить один, вдалеке от людей, как отшельник:  
а) да б) иногда с) нет.
3. Если бы я говорил, что небо «находится снизу», и что зимой «жарко», я должен был бы назвать преступника.  
а) гангстером б) святым с) тучей
4. Когда я ложусь спать, я:  
а) насыпаю мгновенно б) нечто среднее с) засыпаю медленно, с трудом.

В ходе обследования испытуемым раздаются буклеты с текстом опросника и бланки для ответов, в которых нужно последовательно отметить вариант ответа на каждый из вопросов. Полученные результаты выражаются в шкале стенов с минимальным значением в 0 баллов, максимальным – 10 и средним – 5,5 балла (см. параграф 3.2.1). По результатам строится личностный профиль, при интерпретации которого руководствуются степенью выраженности каж-

дого фактора, особенностями их сочетания, а также нормативными данными.

Создание **эмпирических** опросников основано на **критериально-ключевом принципе** (то есть на основе эмпирического обнаружения психологических признаков, позволяющих дифференцировать релевантные критериальные группы от контрольных). Так, при разработке MMPI (Minnesota Multiphasic Personality Inventory) из первоначального перечня утверждений в основные клинические шкалы включались только те, которые хорошо дифференцировали испытуемых с тем или иным клиническим диагнозом от контрольной группы здоровых людей. Другой пример – опросники интересов. Критериально-ключевой принцип достаточно эффективен при построении методик для профотбора. В отличие от факторных опросников, создание эмпирических опросников осуществляется путём поиска заданий, дискриминирующих контрольные группы испытуемых, подобранные на основе критерия, имеющего отношение к тестируемому поведению или свойству личности.

Недостаток эмпирических опросников заключается в дифференциации релевантных групп между собой. Так, в клинической психологии многое зависит от окончательного диагноза, который может быть по-разному поставлен различными специалистами. Эмпирические опросники требуют тщательной психометрической проверки.

**Основные проблемы** использования личностных опросников связаны с возможностью фальсификации ответов и со снижением достоверности вследствие установки испытуемого и различий в понимании вопросов обследуемыми.

Фальсификация ответов типична не для всех диагностических ситуаций. Имеются данные, свидетельствующие о том, что в большинстве случаев испытуемые стремятся быть искренними. Вместе с тем, при проведении некоторых обследований (например, в случае психологической экспертизы) вероятность фальсификации существенна и об этом нужно помнить.

Установки испытуемого проявляются в тенденции выбора социально желательной альтернативы, что может быть связано со стремлением выглядеть лучше. Вместе с тем, по мнению многих исследователей, сила социально желательного ответа связана с более общей потребностью индивида в самозащите, социальном одобрении, в желании избежать критики. С другой стороны, индивид с установкой выбора неблагоприятных ответов может испытывать потребность в чьём-либо внимании, симпатии или помощи.

На достоверность вопросов существенное влияние оказывает интеллектуальная оценка вопросов обследуемыми. Неясность, трудность вопросов снижает надёжность опросника. Текст вопросов должен быть простым и понятным для всех опрашиваемых. Сложные, трудные для понимания и многозначные слова должны исключаться. Каждый вопрос должен содержать только одну мысль и не допускать многозначной интерпретации. С другой стороны, опасны и наводящие вопросы, внушающие определённый ответ. Вопросы должны быть сформулированы так, чтобы их содержание было беспристрастным и социально нейтральным.

Для повышения надёжности опросников используются следующие пути: многократное дублирование вопросов, введение «шкалы лжи», отказ от прямых вопросов.

При анализе результатов диагностики следует также помнить слова С.Л.Рубинштейна о том, что диагностические показатели, получаемые с помощью личностных опросников, имеют «вероятностно-ориентирующие значения», их истинная природа должна быть выявлена путём сопоставления с соответствующими объективными данными. Иными словами, в ответах индивида на вопросы проявляются и свойства нервной системы, и особенности темперамента, и интеллекта, и характерологические особенности. Поэтому попытки объяснить результаты ответов на вопросы чем-то одним некорректны. Впрочем, это ещё раз подтверждает то правило, что диагностика должна быть комплексной и включать как строгоформализованные методики, так и малоформализованные методы оценки (наблюдение, беседа).

### 3.1.3 Психосемантические методики.

Психосемантика (от греч. семантикос – обозначающий) – область психологии, изучающая генезис, строение и функционирование индивидуальной системы значений, опосредствующей процессы восприятия, мышления, памяти, принятия решений и т.п. Психосемантика исследует различные формы существования значений в индивидуальном сознании, анализирует влияние мотивационных факторов и эмоциональных состояний субъекта на формирующуюся у него систему значений. Основным методом экспериментальной психосемантики служит построение субъективных семантических пространств, являющихся модельным представлением категориальных структур индивидуального сознания [76].

Диагностика личности с помощью психосемантических методик представляет собой самоописание (самохарактеристику) субъекта или других лиц с помощью системы значений, отражающих конкретные черты личности. Информация, получаемая с помощью психосемантических методов, опирается на субъективные оценки. А.Анастази называет эти методы феноменологическими (в феноменологической психологии основное внимание уделяется тому, как события воспринимаются индивидом). Психосемантические методики являются идеографическими и применяются для обследования индивидуальных случаев.

Психосемантические методики делятся на две группы:

- списки дескрипторов (описаний черт), с помощью которых испытуемый должен описать себя или других лиц (методика Т.Лири);
- шкальные техники, предназначенные для оценивания себя или других по определённым шкалам, полюса которых обозначены противоположными по значению терминами (семантический дифференциал, Q-сортировка).

Одна из наиболее распространённых методик шкалирования, получившая название «семантический дифференциал» разработана Ч.Осгудом. Применяется в первую очередь для измерения личностных установок и отношений к определённому кругу объектов и

представляет собой стандартизированную методику, в которой испытуемому предъявляется несколько семибалльных шкал, концы которых заданы прилагательными, имеющими противоположные значения (табл. 3.1). Семь градаций могут быть обозначены вербально, численно или графически (в виде отрезков).

**Таблица 3.1 Образец задания методики «Семантический дифференциал»**

	ОТЕЦ	
Хороший	3 2 1 0 1 2 3	плохой
жестокий	3 2 1 0 1 2 3	добрый
серьёзный	3 2 1 0 1 2 3	легкомысленный
замкнутый	3 2 1 0 1 2 3	раскованный
сильный	3 2 1 0 1 2 3	слабый

Каждый объект (понятие) должен быть оценен не менее чем по 15 таким биполярным шкалам. Если соединить отрезками все сделанные отметки, то на бланке отразится субъективный семантический профиль прошкалированного объекта.

При обработке результатов используют два подхода: анализ семантического профиля; построение семантического пространства. Анализ семантических профилей осуществляется через сопоставление числовых оценок всех объектов по всем шкалам. При построении семантического пространства принимают во внимание проведённые Осгудом исследования исходной серии из 50 шкал. Применив корреляционный и факторный анализ он выявил три основных фактора:

**фактор оценки** с высокими нагрузками по таким шкалам, как хороший–плохой, полезный–бесполезный, чистый–грязный;

**фактор силы**, явно проявляемый в шкалах сильный–слабый, большой–маленький, тяжёлый–лёгкий;

**фактор активности**, выявленный в таких шкалах, как активный–пассивный, быстрый–медленный, острый–тупой.

Результаты исследований Ч.Осгуда и других психологов из разных стран мира показали, что система факторов Оценка–Сила–Активность является универсальной (выявляется при описании



различных понятий из разных предметных областей) и устойчивой (воспроизводится в разных языковых культурах).

Семантический дифференциал применяется для решения разных практических проблем: в клинике, в профессиональном консультировании, в маркетинговых исследованиях, для выявления разнообразных социальных установок, для исследования отношения к изучаемым учебным предметам, к товарам и торговым маркам и т.д.

Основная сложность разработки психосемантических методик заключается в подборе и унификации значений слов и выражений, включаемых в состав списков черт, в неизбежности их субъективной интерпретации разными индивидами, когда даже незначительное смещение акцента в ту или иную сторону может привести к искажениям результата.

Этого недостатка лишён разработанный Дж. Келли семантический подход к диагностике личности, получивший название **репертуарных личностных решёток**. При его использовании испытуемый сам конструирует свой индивидуальный опыт, выявляет те субъективные шкалы, которыми он пользуется при восприятии, понимании и оценке окружающего мира [75].

В процессе диагностирования выстраивается матрица, заполняемая либо психодиагностом, либо самим испытуемым.

Столбцы матрицы, называемые элементами, представляют собой группу объектов, в отношении которых выставляются оценки испытуемого. В качестве объектов могут выступать различные предметы, явления, отношения, люди, понятия, ситуации, звуки, цвета, профессии, учебные дисциплины и т.д., т.е. всё то, что психодиагносту представляется существенным для выявления тех или иных отношений или оценок человека. Определение «репертуарная» означает, что элементы в решётку подбираются по определённым правилам (они должны составлять какую-либо область исследования и быть связаны между собой осмысленным образом подобно репертуару ролей в пьесе).

Строками матрицы являются конструкты – субъективные шкалы человека, те мерки, с которыми он подходит к построению образа

окружающего мира. По определению Келли конструкт – это биполярный признак, противоположные отношения или способы поведения (хороший–плохой, настоящее–прошлое, работодатель–работник). Для выявления конструктов используется метод триад. На каждом этапе, получив три карточки с элементами, испытуемый должен назвать какое-нибудь важное качество, по которому два элемента сходны между собой и этим отличаются от третьего. Это качество является одним полюсом конструкта. После этого испытуемого просят назвать, в чём конкретно состоит отличие третьего элемента. Ответ на этот вопрос представляет собой противоположный полюс конструкта. В процессе диагностирования каждый элемент оценивается по каждому конструкту.

Репертуарные решётки могут применяться для проведения психологической оценки в самых различных сферах практики. Их применение требует высокой квалификации.

### **3.1.4 Методики проективной техники**

Методики проективной техники разрабатываются в рамках проективного диагностического подхода. Характерными признаками проективных методик являются:

- использование стимулов, которые испытуемый должен конструировать, дополнять, интерпретировать: например, в тесте тематической апперцепции (ТАТ) испытуемый должен интерпретировать содержание сюжетных картинок (рис. 3.4), в методике Розенцвейга (рис.3.5) – завершать высказывания одного из лиц на сюжетной картинке, в тесте Роршаха (рис. 2.2) – давать толкование неопределённых очертаний чернильных пятен;
- наличие заданий по самостоятельному созданию продукта (рассказа, рисунка, постройки из кубиков).

У такого рода заданий не существует правильного ответа, можно лишь судить о релевантности ответа содержанию задания. Характер ответов испытуемого определяется особенностями его личности, которые «проецируются» в его ответах. Цель обследования в проективных методиках для испытуемого замаскирована, что

уменьшает его возможности произвести желательное впечатление о себе. Эти методики носят в основном индивидуальный характер и в большей своей части это предметные или бланковые методики.



**Рис. 3.4 Образец задания методики ТАТ**

Как правило, проективные методики не обладают очевидной валидностью, то есть цель обследования замаскирована от испытуемого, которому поэтому сложнее осознанно повлиять на результат психодиагностики (фальсифицировать данные). В отличие от опросников проективные методики характеризуются глобальным подходом к оценке личности, а не измерением отдельных черт. Считается, что проективные методики наиболее эффективны при выявлении скрытых, латентных или неосознаваемых сторон личности. Чем менее структурирован стимульный материал, тем он более чувствителен, так как менее структурированные и неоднозначные стимулы гораздо реже вызывают защитные реакции у обследуемого.



**Рис. 3.5 Образец задания методики Розенцвейга**

Исторически сложилось так, что проективные методики разрабатывались и апробировались в условиях психиатрических лечебниц. Методологической основой проективных методик являются традиционные и современные психоаналитические теории. В то же время предпринимались попытки заложить фундамент под проективные методики со стороны бихевиоризма и перцептуализма (поведение человека в каждый момент определяется текущим восприятием). В то же время, например, А.Анастаси считает, что «нет необходимости оценивать конкретные методики в свете теоретического подхода к их созданию или их исторических корней. Методика может оказаться практически полезной или эмпирически ценной по другим причинам, нежели те, которые первоначально приводились для оправдания её введения» [4].

Термин «проекция» впервые был введен З.Фрейдом. Проекция (наряду с вытеснением, рационализацией, сублимацией и т.д.) является одним из защитных механизмов. «Враждебность, о которой ничего не знаешь и также впредь не хочешь знать, переносится из

внутреннего восприятия во внешний мир и при этом отнимается от самого себя и приписывается другим» [71]. Иными словами, проекция – один из защитных механизмов, посредством которого внутренние импульсы и чувства, неприемлемые для «Я», приписываются внешнему объекту и тогда проникают в сознание как изменённое восприятие окружающего мира. Долгое время именно такое толкование проекции механически переносилось на сущность проективной диагностики.

Использование термина «проекция» для обозначения особой группы психодиагностических методов принадлежит Лоуренсу К. Франку. Им были выдвинуты три основных принципа, лежащих в основе проективного исследования личности:

1. Направленность на уникальное в структуре или организации личности. В отличие от традиционных психометрических процедур, личность рассматривается как система взаимосвязанных процессов, а не перечень способностей или черт.

2. Личность в проективном подходе изучается как относительно устойчивая система динамических процессов, организованных на основе потребностей, эмоций и индивидуального опыта.

3. Эта система основных динамических процессов постоянно, активно действует на протяжении жизни индивида, «формируя, направляя, искажая, изменяя и переиначивая каждую ситуацию в систему внутреннего мира индивида». Каждое новое действие, каждое эмоциональное проявление индивида, его восприятия, чувства, высказывания, двигательные акты несут на себе отпечаток личности.

Поскольку стимулы в проективных методиках допускают различную интерпретацию, они приобретают смысл в связи с личностным значением, придаваемым им испытуемыми.

Общими для всех проективных методик являются следующие признаки (по Л.Франку):

- неопределённость, неоднозначность используемых стимулов;
- отсутствие ограничений в выборе ответа;

- отсутствие оценки ответов испытуемых как «правильных» или «ошибочных».

Проективные методики получили наибольшую популярность, начиная с середины XX века и до сих пор занимают лидирующее положение в зарубежных исследованиях личности, прежде всего, в области клинической психодиагностики. Вместе с тем, многие специалисты в области проективной диагностики отмечают необходимость сочетать диагностические данные, полученные с помощью проективных методик, с анкетными данными испытуемых, результатами наблюдений, а также применять наряду с проективными методиками личностные опросники. Дополнительная информация может существенно повысить точность интерпретации данных, полученных с помощью проективных методик.

### **3.1.5 Психофизиологические методики**

Этот вид методик предназначен для диагностики природных особенностей человека, обусловленных основными свойствами его нервной системы. Эти методики лишены оценочного подхода к человеку. Проводя тесты интеллекта или способностей, психолог даёт в конечном счёте оценочное заключение, есть оценочный подход и в некоторых личностных тестах, выявляющих наличие или отсутствие общепризнанных общечеловеческих достоинств. Диагностические психофизиологические методики не претендуют на оценку, поскольку нельзя утверждать, какие свойства нервной системы лучше, а какие хуже. В одних обстоятельствах лучше проявят себя люди с одними свойствами нервной системы, в других – с другими.

По своей форме большинство психофизиологических методик являются аппаратными: используются как электроэнцефалографы, так и специальная аппаратура.

В нашей стране под руководством К.М.Гуревича на основе работ Б.М.Теплова и В.Д. Небылицына разработаны бланковые психофизиологические методики [73]. Валидность и надёжность проверялась при сопоставлении с аппаратными аналогами. Диагно-



стируют проявления основных свойств нервной системы (сила, лабильность) в мыслительно-речевой деятельности.

**Лабильность нервной системы** – это скорость возникновения и исчезновения нервных процессов в коре головного мозга в ответ на внешний раздражитель. Проявления лабильности-инертности нервной системы находят своё выражение в скоростных, темповых показателях выполнения разнообразных видов деятельности:

- в динамике речевого выражения мыслей;
- в скорости актуализации имеющихся знаний;
- в скорости реагирования на внешние раздражители;
- в продуктивности деятельности в ситуации, связанной с лимитом времени (скоростные тесты IQ);
- в лёгкости приспособления к новым видам деятельности.

Примером бланковой психофизиологической методики, диагностирующих лабильность нервной системы, является методика «Исполнение инструкций» – модификация субтеста серии «Альфа» (из него взят принцип составления заданий). Испытуемому нужно быстро выполнить действия по инструкции.

Приведём образцы заданий этой методики:

а). На карточке изображены цифры: **1 3 2 4 5 7 4 8**.

На выполнение задания даётся **4 секунды**.

**Инструкция:** «Проведите линию от цифры 1 к цифре 7 так, чтобы она проходила над чётными цифрами и под нечётными. Начали!»

б). На карточке изображена линия \_\_\_\_\_

На выполнение задания даётся **2 секунды**.

**Инструкция:** «Над линией поставьте стрелку, направленную вниз, а под линией – стрелку, направленную вправо. Начали!»

**Сила нервной системы** – это способность нервных клеток больших полушарий коры больших полушарий головного мозга выдерживать более или менее длительное и концентрированное возбуждение, не переходя в тормозное состояние.

В качестве тестового материала для диагностики силы нервной системы используются задания интеллектуальных тестов, напри-

мер «4-й лишний», «Аналогии» и т.п. При этом исследователя интересует не сам по себе уровень сформированности у испытуемых умения совершать подобные мыслительные операции, а динамика их работоспособности во время их операций. Для этой цели можно использовать и различные модификации корректурных проб.

Рассмотрим пример. Испытуемому предлагается набор заданий «4-й лишний» и даётся инструкция: «Три из четырёх слов в каждом задании являются однородными понятиями и могут быть объединены по одному признаку, а одно слово «лишнее», т.е. выпадает из этой классификации. Вы должны назвать его. После этого вам следует немедленно приступить к решению следующего задания. Время на решение каждого задания ограничено, и, если вы не успеваете отыскать лишнее слово, то по команде экспериментатора «Дальше!» должны сразу же приступить к выполнению следующего задания.

Прежде чем приступить к выполнению задания, испытуемый решает 10 заданий вместе с экспериментатором, который даёт необходимые объяснения. Затем испытуемому предлагается выполнить эти тренировочные задания в темпе, соответствующем основному эксперименту.

Время, отведённое на решение каждого задания, равно 10 секундам. Все задания напечатаны на 5 бланках, по 35 заданий на каждом бланке. Показателем динамики работоспособности служит отношение времени выполнения заданий на последнем бланке ко времени выполнения заданий на первом бланке.

### **Контрольные вопросы и задания**

- 1. Каковы преимущества и недостатки индивидуального и группового обследования? Составьте таблицу. При работе с какими группами лиц не допустима только индивидуальная диагностика?*
- 2. Охарактеризуйте способы использования компьютера при проведении психологических исследований.*

3. Дайте характеристику психосемантическому подходу в психодиагностике. Назовите основные факторы семантического пространства. На какие группы делятся психосемантические методики?
4. Опишите метод репертуарных решёток. Является ли он строго-формализованным и почему? Что такое конструкт в методе репертуарных решеток?
5. В чём специфика проективного подхода в психодиагностическом изучении личности?
6. Пользуясь психологическим словарём [76], словарём-справочником по психодиагностике [11] и другой литературой по психодиагностике, заполните таблицы 3.2, 3.3 и 3.4.

**Таблица 3.2 Классификация тестов по содержанию**

Тесты	Предназначение	Специфика	Формы	Примеры
интеллекта				
специальных способностей				
личности				
достижений				

\*) В колонке «Формы» перечисляются те формы тестов, к которым студент может привести пример.

**Таблица 3.3 Классификация опросников**

№	Тип опросника	Содержание	Примеры
1			
...			

**Таблица 3.4 Классификация проективных методик**

№	Тип	Содержание	Примеры
1	Конститутивные		
2	Конструктивные		
3	Интерпретативные		
4	Катартические		
5	Экспрессивные		
6	Импрессивные		
7	Аддитивные		

## **3.2. Метрические характеристики строгоформализованных методик**

Психодиагностические методики являются измерительными инструментами. Для того, чтобы измерить какой-либо мерой величину измеряемого, необходимо, чтобы были соблюдены следующие условия (К.М.Гуревич [17, 19]). Первое условие касается самого измерительного инструмента: он должен иметь на протяжении всего измерения одну и ту же, тождественную самой себе меру. Нельзя измерять инструментом, у которого нет такой тождественной самой себе меры (например, если бы на линейке расстояния от одного деления до другого были неодинаковы на разных участках). Второе условие – нужно, чтобы то, что измеряется, во всех случаях оставалось одним и тем же, т. е. также было тождественно самому себе. Как показывает далее К. М. Гуревич, соблюсти эти два условия в психодиагностических исследованиях крайне затруднительно. Вопрос стабильности измеряемых психических характеристик находится в компетенции соответствующих предметных областей психологической науки. В настоящем параграфе будут рассмотрены требования к психодиагностическим методикам, как профессиональным инструментам измерения индивидуально-психологических характеристик. К числу этих требований относятся стандартизация, надёжность и валидность.

### **3.2.1. Стандартизация**

Стандартизация – это единообразие процедуры проведения и оценки выполнения теста. Она рассматривается в двух планах: как выработка единых требований к процедуре эксперимента и как определение единого критерия оценки результатов (рис. 3.6).

**Стандартизация процедуры проведения** обследования подразумевает единообразие инструкций, бланков и других материалов обследования, способов регистрации результатов, условий проведения обследования.



**Рис. 3.6 Составляющие процесса стандартизации психодиагностической методики**

При проведении психодиагностического обследования нужно соблюдать следующие требования:

- инструкции следует сообщать испытуемым одинаковым образом, лучше письменно; в случае устных указаний они даются в разных группах одними и теми же словами, понятными для всех, в одинаковой манере;
- ни одному испытуемому не следует давать никаких преимуществ перед другими;
- в процессе эксперимента не следует давать отдельным испытуемым дополнительные пояснения;
- эксперимент с разными группами следует проводить по возможности в одно и то же время дня в сходных условиях;
- временные ограничения в выполнении заданий для всех испытуемых должны быть одинаковыми.

Как правило авторы методики приводят в методическом руководстве точные и подробные указания по процедуре проведения обследования. Формулирование таких указаний составляет основ-

ную часть стандартизации новой методики, так как только строгое их соблюдение даёт возможность сравнивать между собой показатели, полученные разными испытуемыми.

**Стандартизация критериев выполнения** теста осуществляется путём её проведения на большой представительной выборке испытуемых, которая ничем не отличается от той, для которых данная методика предназначена.

Для этих целей должна быть сформирована **выборка** стандартизации, представляющая собой меньшую по размеру точную модель генеральной совокупности. Если выборка стандартизации сформирована с соблюдением всех правил, изложенных ниже, то результаты её обследования можно полноправно переносить на всю генеральную совокупность.

Формирование выборки стандартизации следует производить с применением **метода стратифицированной случайной выборки**. Для этого элементы генеральной совокупности разбиваются на страты (группы) в соответствии с некоторыми характеристиками и в дальнейшем производится случайная выборка из каждой страты. На этой группе испытуемых высчитываются нормы, указывающие на высокий, средний и низкий уровень выполнения задания. Впоследствии в результате тестирования определяется положение каждого испытуемого относительно выборки стандартизации. Иначе говоря, выборка стандартизации является **контрольной группой**, с которой впоследствии сопоставляются результаты обследованных лиц.

При формировании выборки следует учитывать её объём и репрезентативность. Выборка должна быть достаточно большой для обеспечения такой величины стандартной погрешности измерений, которой можно было бы пренебречь. С другой стороны, выборка должна точно отражать категорию лиц, для которых предназначен тест.

Для простого уменьшения значения стандартной погрешности вполне адекватной будет выборка из 500 испытуемых. Однако репрезентативность выборки не зависит от её объёма. При получении



нормы для общей популяции детей школьного возраста необходима выборка около 10 000 испытуемых. А выборка из укротителей львов (пример приведён П.Клайном [29]), конечно, не должна быть столь большой. Таким образом, нельзя сделать никакого утверждения относительно объёма выборки безотносительно той популяции (категории лиц), из которой она подбирается. Репрезентативность выборки является более важной, чем её размер. Маленькая, но репрезентативная нормативная выборка будет предпочтительнее, чем большая, но неравномерно представленная.

Рассмотрим два примера.

**Пример 1.** При формировании выборки стандартизации для теста интеллекта Лордж-Торндайка было выделено 12 возрастных групп: от 6 до 17 лет. Общее количество испытуемых 136 000, т.е. более чем 11 000 человек на каждую возрастную группу. Для обеспечения репрезентативности была взята стратифицированная (разделённая на группы, классы) выборка из разных социальных групп, а основой для стратификации послужили факторы, которые разработчики сочли наиболее связанными с уровнем интеллекта:

- а) процент грамотного взрослого населения;
- б) пропорция профессиональных рабочих в популяции;
- в) процент домовладельцев;
- г) среднее количество снимающих жильё.

Каждая социальная группа была классифицирована по этим 4-м переменным на очень высокий, высокий, средний, низкий и очень низкий уровни. Затем были протестированы все учащиеся каждого уровня в каждой из социальных групп, и для каждого из этих уровней в каждой группе были получены отдельные нормы. Фактически, общее количество испытуемых было вычислено для 44-х социальных групп США.

Исследование Лордж-Торндайка показывает, каким должен быть полный объём выборки, необходимый для нормальной нормализации, и громадный объём административных расходов, требующихся для проведения такого обширного тестирования. Под-

бор испытуемых в рамках стратификационных групп должен быть случайным.

**Пример 2.** При формировании выборки стандартизации для 16ФЛО Р. Кеттелла были определены следующие основания для стратификации: социальное положение, географическое положение, возраст, горожанин или сельский житель. Выборка стандартизации отражала пропорции этих категорий во всём населении. Общий объём выборки 977 человек. Относительно небольшой её объём связан с проблемой нежелания сотрудничать с исследователями взрослых испытуемых, не являющихся военнослужащими, студентами или заключёнными.

**Правила формирования выборок** для специальных групп (П.Клайн [29]):

1. Найти наиболее важные переменные, релевантные (значимые, существенные) для данных групп, и использовать их как основание для стратификации.
2. Подобрать настолько большую выборку, насколько возможно: как минимум 300 человек.
3. Если использовались небольшие выборки, акцентировать на этом внимание пользователей теста.

**Стандартизация оценки выполнения** применяется в основном для тестов и опросников и включает определение нормативных данных и построение шкалы стандартных показателей.

**Нормы** – это статистические показатели выборки стандартизации (среднее значение и стандартное отклонение), позволяющие охарактеризовать положение обследованного индивидуума относительно выборки стандартизации. Рассчитанные на основе обследования выборки стандартизации среднее значение (обозначается латинской буквой  $M$  или греческой буквой  $\mu$  – мю) и стандартное отклонение (обозначается греческой буквой  $\sigma$  – сигма) являются основой для разработки шкалы оценок в стандартных показателях.

Тестовые нормы определяются отдельно для каждой группы испытуемых (в тестах интеллекта, как правило, группы возрастные;

также возможно представление тестовых норм для испытуемых разного пола, уровня образования, профессии и т.д.) Рестандартизация и определение норм для новой социоэтнической популяции является одним из основных и обязательных элементов адаптации теста. Многие психодиагностические методики требуют периодического (один раз в 5-7 лет) пересмотра норм.

После того, как получены результаты обследования выборки стандартизации, производится их проверка на нормальность. От этого зависит определение шкалы оценок. Если распределение отличается от нормального, следует пересмотреть содержание заданий теста (например, их трудность, либо время выполнения). Если распределение нормальное, то можно использовать один из типов стандартных показателей.

Все стандартные показатели рассчитываются на основе параметров нормального распределения: среднего значения  $\mu$  и стандартного отклонения  $\sigma$ , которые вычисляются по формулам (3.1) и (3.2) соответственно.

$$\mu = \frac{\sum_{i=1}^n x_i}{n} \quad (3.1)$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}} \quad (3.2)$$

Единичной нормальной кривой называется нормальная кривая с  $\mu = 0$  и  $\sigma = 1$ . Для любой нормальной кривой (рис. 3.7)

68.27% площади лежит в интервале  $[\mu - \sigma, \mu + \sigma]$ .

95.45% – в интервале  $[\mu - 2\sigma, \mu + 2\sigma]$ .

99,73% – в интервале  $[\mu - 3\sigma, \mu + 3\sigma]$ .

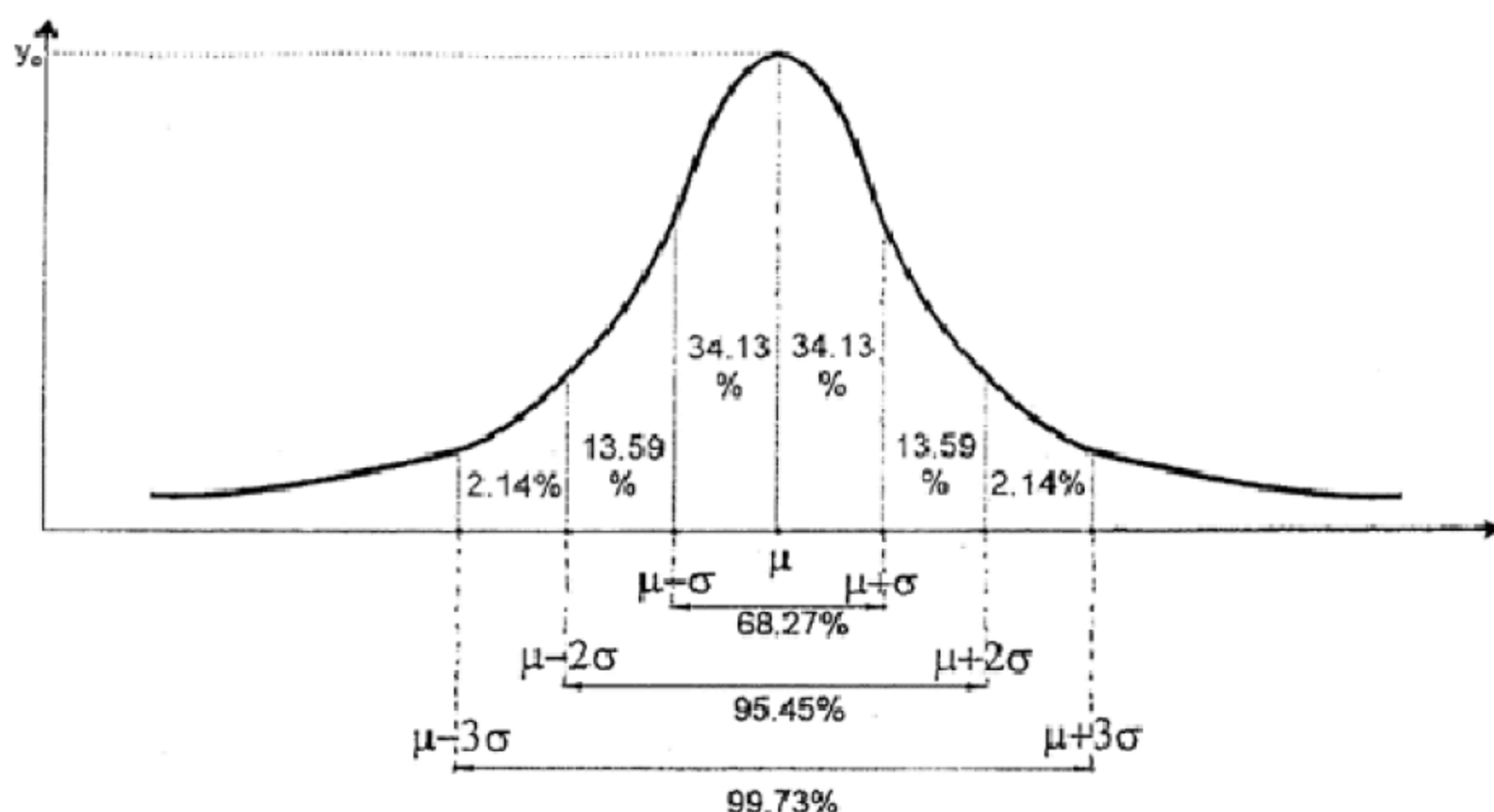
При разработке шкалы стандартных показателей используют следующие правила:

1. Результаты, которые лежат в диапазоне  $[\mu - \sigma, \mu + \sigma]$  (эти результаты показали 68.27 % испытуемых выборки стандартизации) считаются **нормальными** (средними) для данной генеральной совокупности.

2. Результаты в диапазоне  $[\mu + \sigma, \mu + 2\sigma]$  оцениваются, как результаты «**выше среднего**», а результаты в диапазоне  $[\mu - 2\sigma, \mu - \sigma]$

оцениваются, как результаты «**ниже среднего**» для данной генеральной совокупности. Такие результаты показали по 13,6% испытуемых выборки стандартизации.

3. Результаты в диапазонах  $[\mu - 3\sigma, \mu - 2\sigma]$  и  $[\mu + 2\sigma, \mu + 3\sigma]$  оцениваются как «**низкие**» и «**высокие**». Такие результаты показали по 2,14% испытуемых выборки стандартизации.



**Рис. 3.7. Нормальное распределение**

Следует подчеркнуть, что в ходе практического использования методики оценка результатов каждого испытуемого (т.е. её сопоставление с тестовыми нормами), имеет смысл в том и только том случае, если этот **испытуемый относится к той же генеральной совокупности, что и выборка стандартизации**. В этом, и только этом случае сырые баллы переводятся в стандартные показатели, к которым относятся Z-показатели, T-показатели, станы, станайны.

### **Z-показатели**

Z-показатель — выражение шкальных показателей в единицах стандартного отклонения:

$$Z = \frac{x - \bar{x}}{\sigma} \quad (3.3)$$

где  $x$  – «сырой» показатель.

При выполнении условия нормального распределения исходных результатов, шкалы, основанные на  $Z$ -показателях, являются интервальными.

$Z$ -показатели не всегда удобны для окончательного представления данных, так как среднее значение в полученной шкале равно 0, стандартное отклонение – 1, а наивысшее значение – 3 и при этом существуют отрицательные значения. Этот факт создаёт дополнительные проблемы, особенно при работе с клиентами. Поэтому разработчики тестов применяют дальнейшее преобразование  $Z$ -показателей, которое заключается в следующем:

$$Z_t = A + B \cdot Z, \text{ где} \quad (3.4)$$

$Z_t$  – преобразованный  $Z$ -показатель;

$A$  – среднее значение полученного распределения;

$B$  – стандартное отклонение преобразованного распределения.

В случае  $A=50$  и  $B=10$  получается **Т-показатель**, который часто используется разработчиками различных тестов. Так, в ММРІ все сырые показатели переводятся именно в показатели Т-шкалы.

Аналогично, в шкале Стэнфорд-Бине  $A=100$ ,  $B=16$ .

В таблице 3.5 представлены нормы теста Равена в Т-баллах, полученные нами для учащихся начальной школы на выборке 280 учащихся г.Обнинска в 2004 году.

**Таблица 3.5 Нормы теста Равена в Т-баллах для учащихся 1-3 классов**

Т-шкала	1 класс	2 класс	3 класс
30 и менее	8 и менее	10 и менее	21 и менее
31-40	9-17	11-20	22-30
41-50	18-27	21-30	31-38
51-60	28-36	31-40	39-47
61-70	37-46	41-50	48-56
71-80	47-55	51-59	57-60

Если необходимо сопоставить сырые результаты по двум тестам и данные по обоим тестам распределены нормально, то следует привести данные к стандартным нормальным показателям с одинаковыми  $A$  и  $B$ , вычислить, каким сырым значениям соответствуют показатели нормализованной шкалы, и после этого сопоставить значения.

Следующие две шкалы стандартных показателей основаны на разделении диапазона оценок на интервалы, размер которых не привязан к стандартному отклонению, а количество интервалов задаётся исследователем. Эти шкалы получили название:

- станайны (от англ. stanine – standard nine);
- стены (от англ. sten – standard ten).

Заметим, что в отечественной психологической литературе иногда встречается иное написание этих шкал: стенайны, стэны. Это связано с различиями в написании и произношении английских слов stanine и sten. Мы будем использовать тот вариант русского язычного написания, который в последнее время закрепился в научном психологическом лексиконе.

### **Станайны.**

Станайн – это нормализованный стандартный показатель на одномерной шкале ранжирования от низшего значения, равного 1 до высшего, равного 9. Этот показатель разбивает нормальное распределение на девять интервалов-категорий. В шкале станайнов  $\mu = 5$  и  $\sigma \approx 2$ . Шкала станайнов сочетает в себе достоинства стандартных шкальных показателей и простоту процентилей. Для преобразования первичных показателей в станайны результаты обследования выборки стандартизации упорядочивают по возрастанию, и из них образуют группы с числом лиц, пропорциональным определённым частотам оценок в нормальном распределении тестовых результатов (табл.3.6).

Отметим, что процентные соотношения подобраны так, что шкала станайнов является интервальной.



Пусть в выборке стандартизации 500 человек, тогда по 20 (4%) испытуемых, имеющих самые низкие и самые высокие показатели будут отнесены к 1 и 9 станайнам соответственно, и так далее.

**Таблица 3.6 Соответствие станайнов процентным долям выборки стандартизации**

Процент обследованных в выборке стандартизации	4	7	12	17	20	17	12	7	4
Станайн	1	2	3	4	5	6	7	8	9

### Стены

При трансформации оценок в шкалу стенов (standard ten) проводится аналогичная процедура с 10 интервалами. Этот показатель разбивает нормальное распределение на десять интервалов-категорий. В шкале стенов  $\mu = 5.5$  и  $\sigma \approx 2$ . Для преобразования первичных показателей в стены результаты обследования выборки стандартизации упорядочивают по возрастанию, и из них образуют группы, пропорциональные по объёму определённым частотам оценок в нормальном распределении результатов (табл.3.7).

**Таблица 3.7 Соответствие стен ов процентным долям выборки стандартизации**

Процент обследованных в выборке стандартизации	0,8	2,8	7,9	15,9	22,6	22,6	15,9	7,9	2,8	0,8
Стен	1	2	3	4	5	6	7	8	9	10

Отметим, что процентные соотношения здесь также подобраны таким образом, что шкала стенов является интервальной: диапазон  $[\mu - 3\sigma, \mu + 3\sigma]$  разделён на 10 равных частей с шагом  $0,6\sigma$ , а затем вычислены процентные доли испытуемых в каждом интервале.

Возьмём рассмотренную выше выборку стандартизации в объёме 500 человек. После упорядочивания испытуемых по результатам

обследования по 4 (0,8 %) испытуемых, имеющих самые низкие и самые высокие показатели, будут отнесены к 1 и 10 стенам соответственно, по 14 (2,8%) – к 2 и 9 стенам, по 40 (7,9%) – к 3 и 8 стенам, по 80 – к 4 и 7 стенам, и, наконец, по 112 человек – к 5 и 6 стенам. Диапазон сырых баллов респондентов выборки стандартизации, отнесённых к каждому стен у, является основой перевода сырых оценок в стены.

В таблице 3.8 представлен пример таблицы перевода сырых баллов по опроснику Р.Кеттелла в стены для испытуемых мужского пола в возрасте от 25 до 33 лет [84].

**Таблица 3.8 Таблица перевода сырых баллов 16ФЛО в стены**

Стэны	1	2	3	4	5	6	7	8	9	10
Факторы										
A	0-4	5-6	7	8-9	10-11	12	13-14	15-16	17-18	19-20
B	0-3	4	5	6	7	8	9	10	11	12-13
C	0-5	6-7	8-9	10-11	12-13	14-15	16-18	19-20	21-22	23-26
E	0-3	4-5	6-8	9-10	11-13	14-15	16-18	19-21	22-23	24-26
F	0-4	5-6	7-8	9-10	11-12	13-14	15-16	17-18	19-20	21-26
G	0-4	5-7	8-9	10	11-13	14	15-16	17	18-19	20
H	0-1	2-4	5	6-8	9-10	11-15	16-17	18-19	20-21	22-26
I	0-1	2-3	4	5-6	7	8-9	10-11	12-14	15-16	17-20
L	0-4	5	6-7	8-9	10	11-12	13	14-15	16	17-20
M	0-5	6-7	8	9-10	11	12-13	14	15-16	17	18-26
N	0-4	5	6-7	8	9	10-11	12	13	14-15	16-20
O	0-1	2-3	4-5	6-7	8-9	10-11	12-13	14-15	16-17	18-26
Q1	0-2	3-4	5	6	7-8	9-10	11-12	13	14-15	16-20
Q2	0-4	5-6	7	8-9	10	11-12	13	14-15	16-17	18-20
Q3	0-6	7	8-9	10	11-12	13-14	15	16-17	18	19-20
Q4	0-2	3-5	6-7	8-9	10-11	12-14	15-16	17-18	19-21	22-26

### Процентили

При проведении рестандартизации тестов, а также в некоторых других случаях, когда, с одной стороны, не удаётся получить нормального распределения результатов обследования в выборке стандартизации, а, с другой стороны, изменение заданий является нежелательным, для представления тестовых норм используют

процентили (см. главу 6). Шкала процентилей не является интервальной (является порядковой).

Ранг показателя в процентилях определяется процентным соотношением в выборке стандартизации тех испытуемых, которые получили более низкий показатель. Так, значение 15-го процентиля ( $P_{15}$ ) означает, что 15% из популяции имеют показатели ниже данного. В качестве примера приведем данные стандартизации теста Равена на выборке взрослых жителей Великобритании (табл.3.9). Из таблицы видно, что 75% выборки набрало не более 57 баллов и всего лишь 10% – меньше 44.

**Таблица 3.9 Нормы теста Равена в процентилях**

Процентиль	5	10	25	50	75	90	95
Число баллов	39	44	49	54	57	58	59

### **Представление норм**

При описании норм для тестов должна быть представлена следующая информация:

- 1) данные о выборке стандартизации (её объём и основные характеристики испытуемых: возраст, пол, образование, социальная принадлежность и т.п. уровни стратификации);
- 2) тип нормативных показателей;
- 3) для каждой нормативной группы должны быть указаны «сырые» показатели сравнительно со стандартизированными показателями.

### **3.2.2. Трудность заданий теста**

**Индекс трудности** задания вычисляется по формуле:

$$I = \left(1 - \frac{N_1}{N}\right), \quad (3.5)$$

где  $N_1$  – число испытуемых, правильно выполнивших задание,  $N$  – общее число испытуемых.

Задание считается оптимальным по трудности для данной выборки испытуемых, если индекс трудности близок к 0,5. Отметим, что вычисление индекса трудности применяется для интеллектуальных тестов и тестов достижений.

### 3.2.3. Дискриминативность заданий

Показателем дискриминативности задания служит мера соответствия его успешного решения успешному решению всех заданий теста. Он называется коэффициентом дискриминации и вычисляется как бисериальный (рангово-бисериальный) коэффициент корреляции между результатом выполнения данного задания (дихотомическая шкала) и результатом выполнения всего теста (порядковая или интервальная шкала) у выборки испытуемых.

Если шкала измерения – интервальная, то используется бисериальный коэффициент корреляции:

$$r_{pb} = \frac{\bar{X}_1 - \bar{X}}{\sigma} \sqrt{\frac{N_1}{N - N_1}}, \quad (3.6)$$

где  $\bar{X}_1$  – среднее арифметическое оценок по тесту у тех испытуемых, которые верно выполнили задание;  $\bar{X}$  – среднее арифметическое всех индивидуальных оценок по тесту;  $\sigma$  – стандартное отклонение оценок по тесту в выборке.

Коэффициент дискриминации может принимать значения от  $-1$  до  $+1$ . Высокий положительный  $r_{pb}$  свидетельствует об эффективности деления испытуемых, чем ближе значение коэффициента дискриминации к  $1$ , тем более соответствует данная задача всему тесту. Отрицательное значение коэффициента отдельного задания теста свидетельствуют о несоответствии результатов выполнения этого задания суммарному показателю по тесту. Такие задания должны быть исключены из теста. Отсутствие зависимости между отдельным заданием и всем тестом (значение  $r_{pb}$  по модулю меньше критического) может говорить либо о высокой трудности зада-

ния, либо о том, что оно слишком лёгкое для данной выборки испытуемых.

В тех случаях, когда одна переменная (Y) измеряется в дихотомической шкале, а другая (X) – в порядковой шкале, для определения дискриминативности задания используется рангово-бисериальный коэффициент корреляции  $r_{rb}$ , который вычисляется по формуле:

$$r_{rb} = \frac{2 \cdot |\bar{X}_1 - \bar{X}_0|}{N}, \quad (3.7)$$

где  $\bar{X}_1$  – среднее по тем элементам переменной X, которым соответствует значение 1 переменной Y;  $\bar{X}_0$  – среднее по тем элементам переменной X, которым соответствует значение 0 переменной Y.

### 3.2.4 Надёжность

Надёжность – это психометрическая характеристика строго-формализованной психодиагностической методики, отражающая точность психодиагностических измерений, а также устойчивость результатов теста к действию посторонних случайных факторов. Надёжность теста – это характеристика того, в какой степени выявленные у испытуемых различия по тестовым результатам являются отражением действительных различий в измеряемых свойствах и в какой мере они могут быть приписаны случайным ошибкам.

В психометрике термин надёжности имеет два значения: надёжность по внутренней согласованности и надёжность ретестовая.

#### **А. Надёжность по внутренней согласованности.**

1. *Надёжность эквивалентных половин теста.* Вычисляется корреляция результатов по двум эквивалентным частям теста, полученных после процедуры расщепления (например, на чётные и нечётные задания в тесте Равена). Этот тип коэффициента надёжности называется коэффициентом внутренней согласованности.

2. *Надёжность по внутренней согласованности* – это оценка степени выраженности интеркорреляционных связей между зада-

ниями, составляющими тест. Этот показатель напрямую связан с величиной теста и вычисляется с помощью коэффициента Кронбаха:

$$\alpha = \frac{k}{k-1} \left[ 1 - \frac{\sum_{j=1}^k S_j^2}{S_x^2} \right], \quad (3.8)$$

где  $k$  – число заданий в тесте,  $S_j^2$  – дисперсия по  $j$ -му пункту теста в выборке,  $S_x^2$  – дисперсия суммарных баллов по всему тесту.

#### Б. Надёжность ретестовая.

1. *Ретестовая надёжность* – степень согласованности результатов теста, получаемых при первичном и вторичном его применении на одних и тех же испытуемых, при условии, что испытуемые не изменились. Этот тип надёжности характеризует устойчивость методики. Интервал должен быть не менее месяца. Вычисляется с помощью коэффициента корреляции показателей при первичном и повторном обследовании (ретестовая надёжность считается удовлетворительной при величине коэффициента корреляции не менее 0.7).

2. *Надёжность взаимозаменяемых форм* (при использовании параллельных наборов заданий). В этом случае предварительно требуется доказать, что используемые параллельные формы одной методики являются действительно эквивалентными. Значение коэффициента надёжности определяется по коэффициенту корреляции, полученному между результатами по двум формам теста.

Проверка психодиагностических методик на надёжность должна производиться на достаточно больших (200 и более) и репрезентативных выборках испытуемых.

Основные источники неудовлетворительной надёжности:

1. Наличие в тестах заданий, которые допускают неоднозначное толкование.



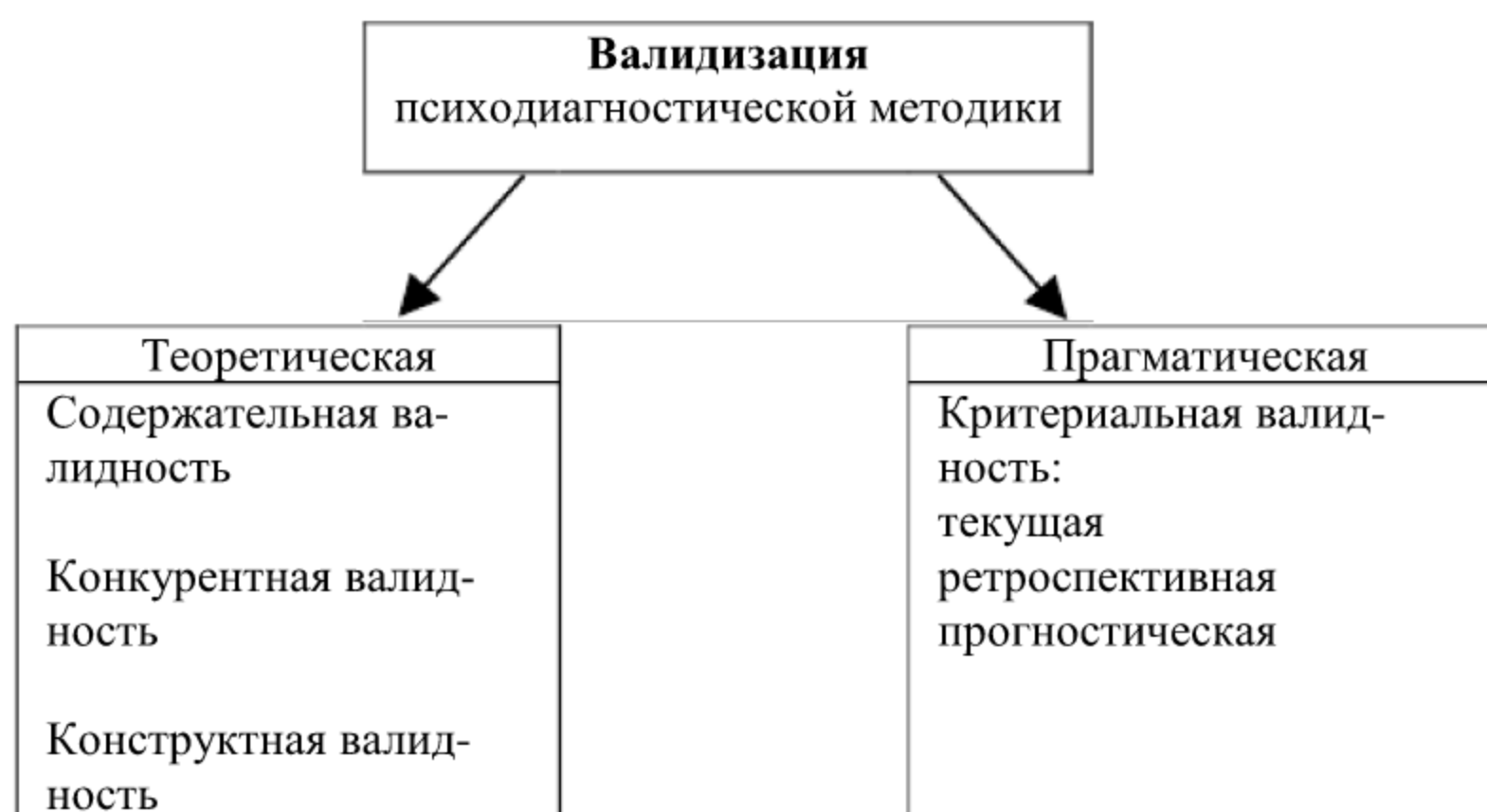
2. Нарушение инструкции и технологии проведения влияет как на трудность заданий, так и на отношение испытуемых к процедуре обследования, что, в свою очередь, сильно влияет на результат.

3. Низкая надёжность психодиагностической методики может быть следствием неустойчивости психологического конструкта, который измеряется при помощи этой методики. Так, например, низкая надёжность характерна для многих тестов креативности [4].

### 3.2.5. Валидность

Тест называется валидным, если он измеряет то, для измерения чего он предназначен. Если показатели надёжности и дискриминативности могут быть вычислены определённо и однозначно, то установление валидности – гораздо более разносторонняя и менее однозначная процедура.

**Валидность** – комплексная характеристика, включающая сведения о том, **пригодна** ли методика для измерения того, для чего она была создана, и какова её **действенность**, практическая полезность. Соответственно процесс проверки валидности (валидизация) в первом случае называется **теоретической валидизацией**, а во втором – **прагматической валидизацией** (рис. 3.8).



**Рис. 3.8 Валидизация строгоформализованной методики и основные типы валидности**

Провести теоретическую валидизацию методики – это значит доказать, что методика измеряет именно то психическое качество, которое она по замыслу исследователя должна измерять. *Например, если какой-либо тест разрабатывался для того, чтобы диагностировать умственное развитие детей, надо проанализировать, действительно ли он измеряет умственное развитие, а не какие-то другие особенности (например, личностные)* Такая проверка показывает, действительно ли замысел автора и результаты применения методики совпадают.

Теоретическую валидизацию новой методики провести менее сложно, если для изучаемого свойства уже имеется методика с доказанной валидностью. Наличие положительной корреляции между новой и аналогичной, уже проверенной методикой, указывает на то, что разработанная методика измеряет то же психологическое качество, что и эталонная. Если такой путь проверки невозможен, психологический смысл методики может быть раскрыт только следующим путём:

- постепенное накопление разнообразной информации об изучаемом свойстве;
- анализ теоретических предпосылок и экспериментальных данных;
- значительный опыт работы с методикой.

Прагматическая валидизация подразумевает проверку методики с точки зрения её практической эффективности, значимости, полезности, поскольку диагностической методикой имеет смысл пользоваться только тогда, когда доказано, что измеряемое свойство проявляется в определённых жизненных ситуациях, в определённых видах деятельности.

Для проведения прагматической валидизации методики обычно используется независимый внешний критерий – показатель проявления изучаемого свойства в повседневной жизни. Таким критерием может быть:

- для тестов способностей к обучению, тестов достижений, тестов интеллекта – успеваемость;

- для методик профессиональной направленности – производственные достижения;
- для тестов специальных способностей – эффективность реальной деятельности (рисование и т.п.);
- для тестов личности – субъективные оценки.

Вывод о валидности методики получается в результате сопоставления результатов диагностики с соответствующим критерием.

Существует несколько различных типов валидности, обусловленных особенностями диагностических методик.

Рассмотрим следующие основные типы валидности:

1. Валидность по содержанию.
2. Критериальная валидность.
3. Конкурентная валидность.
4. Конструктная валидность.

### **Содержательная валидность**

Это один из основных типов валидности методики, характеризующий **степень репрезентативности содержания заданий теста измеряемой области психических свойств**. Содержательная валидность отражает адекватность модели тестируемой деятельности, правильность подбора заданий, охватывающих главные аспекты изучаемого феномена. Чаще всего определяется для тестов достижений.

Так, например, в тесты достижений включается не весь материал, который был изучен, а какая-то его небольшая часть. В этом случае нужно быть уверенным, что правильные ответы на немногие вопросы свидетельствуют об усвоении всего материала.

Содержательная валидность является основным типом валидности критериально-ориентированных тестов, также относящихся к тестам достижений. В традиционных тестах, ориентированных на статистические нормы, которые определяются в результате стандартизации, результаты конкретного испытуемого фактически сопоставляются с результатами выборки стандартизации (т.е. оценивается сравнительная успешность выполнения теста). В критериально-ориентированных тестах указываются и определяются цели

обучающего курса, которые закладываются в тест с помощью заданий, измеряющих степень достижения этих целей. Информация, получаемая с помощью этих тестов – усвоил ли отдельный испытуемый изученный материал или нет.

### **Критериальная валидность.**

Это комплекс характеристик, включающий текущую, прогностическую и ретроспективную валидность.

**Текущая валидность** отражает способность теста дифференцировать испытуемых по изучаемому признаку. В качестве таких признаков могут выступать познавательные процессы, способности, личностные характеристики. Определяется через сопоставление результатов тестирования с экспертными оценками по соответствующим критериям.

**Прогностическая валидность** отражает сведения, характеризующие степень обоснованности и статистической надёжности развития исследуемой психологической особенности в будущем. Заключение об этом типе валидности может быть получено путём сравнения тестовых оценок с некоторым критерием, характеризующим измеряемое свойство спустя некоторое время. Например, прогностическая валидность теста может быть продемонстрирована корреляцией показателей теста в возрасте 11 лет с успеваемостью в 16 лет. Многие специалисты по психодиагностике рассматривают прогностическую валидность как наиболее убедительное подтверждение эффективности теста.

**Ретроспективная валидность** определяется на основе критерия, отражающего события или состояние изучаемого качества в прошлом. Например, для проверки того, в какой мере хорошие результаты теста способностей соответствуют успешному обучению, можно сопоставить прошлые оценки успеваемости, прошлые экспертные заключения и т.д. у лиц с высокими и низкими на данный момент диагностическими показателями.

Основная проблема критериальной валидации заключается в подборе адекватного внешнего критерия.

Д.Тиффин и Е.Маккормик выделили следующие четыре типа внешних критериев [Цит. по 18]:

- критерии исполнения (объём выполненной работы, успеваемость, темп роста квалификации);
- субъективные критерии (оценки, полученные с помощью интервью, опросников, анкет);
- физиологические критерии (пульс, давление, симптомы утомления и т.п. –используются при изучении влияния окружающей среды и других ситуационных переменных на организм и психику человека);
- критерии случайностей (используются, когда цель исследования касается проблемы отбора для работы таких лиц, которые менее подвержены несчастным случаям).

М.К.Акимова определяет следующие требования, которым должен удовлетворять внешний критерий [75]:

- **релевантность** (смысловое соответствие и качественная однородность по психологической сущности); в критерии должны быть задействованы именно те особенности индивидуальной психики, которые измеряются и диагностической методикой.
- **контаминация** (свобода от помех) – низкая чувствительность критерия от внешних условий, не связанных с индивидуальными особенностями, измеряемыми методикой, либо оценивать по критерию нужно людей, находящихся в максимально одинаковых условиях.
- **надёжность** – критерий должен отражать постоянство и устойчивость исследуемой функции.

По методу сопоставления результатов теста с внешним критерием определяются ещё два типа валидности: инкрементная валидность и дифференциальная валидность, которые имеют значение для тестов профотбора и уточняют данные о прогностической валидности теста.

Для определения **инкрементной** валидности вычисляется корреляция субтеста с критерием и другими субтестами батареи. Если коэффициент корреляции субтеста с критерием значим и положи-

телен, а корреляция с другими субтестами незначима, то имеет место инкрементная валидность.

**Дифференциальная** валидность имеет место, когда внешний критерий сам является интегральным показателем (например, академическая успеваемость) и результат теста имеет различную корреляцию с различными аспектами данного критерия (различные учебные дисциплины).

### **Конкурентная валидность**

Оценивается по корреляции разработанного теста с другими, валидность которых относительно измеряемого параметра установлена.

Правила изучения конкурентной валидности:

1. Выборка испытуемых должна соответствовать той генеральной совокупности, для которых данный тест предназначен (пол, возраст, образование, социальное положение).
2. Объём выборки – не менее 200 чел.
3. Использовать как можно большее число других валидных тестов данной переменной (например, для тестов интеллекта – различные вербальные и невербальные субтесты).
4. При использовании факторного анализа добиться простой структуры.
5. Психологическое значение полученных результатов следует из информации о корреляциях и факторных нагрузках.

### **Конструктная валидность**

Это один из основных типов валидности, отражающий степень теоретической обоснованности теста. Иными словами, это соответствие результатов тестирования показателям тех психологических качеств (конструктов), которые теоретически связаны с оцениваемым свойством. В качестве конструкта могут выступать познавательные процессы, личностные особенности, общие способности (интеллект, креативность). Процедура установления конструктной валидности сложнее, чем определение содержательной или критериальной валидности, так как проявления таких конструктов в дея-



тельности человека многообразны и неоднозначны с точки зрения их выделения (пример – интеллект).

Проверка конструктивной валидности проводится в три основных этапа:

1) определение некоторой теоретической концепции, которая предположительно объясняет выполнение теста;

2) из теоретической концепции выводятся гипотезы, связанные с тестом:

- относительно переменных, с которыми результат теста должен коррелировать (конкурентная валидность);
- относительно переменных, с которыми результат теста не должен коррелировать;
- указываются группы испытуемых, которые должны получать высокие и низкие оценки по данному тесту;

3) выдвинутые гипотезы подвергаются эмпирической проверке.

Если эмпирические данные подтверждают гипотезу, то тем самым подтверждается психологическая концепция, положенная в основу теста, и способность теста служить инструментом измерения данного конструкта.

В заключение остановимся ещё на одном типе валидности. Это валидность с точки зрения испытуемого, или **очевидная валидность**. Она определяет, насколько адекватно впечатление о предмете измерения, которое формируется у испытуемого при знакомстве с инструкцией и материалом теста, тому, что данный инструмент в действительности выявляет. Очевидная валидность не имеет никакого отношения к истинной валидности (действенности) методики.

Очевидная валидность присуща тесту, если реальные цели тестирования от испытуемого не скрыты, что имеет свои плюсы и минусы. С одной стороны, некоторые испытуемые не включаются в работу с тестами, которым недостаёт очевидной валидности, полагая порой их глупыми и даже оскорбительными. В этих случаях очевидная валидность тестов помогает добиться сотрудничества с испытуемыми. С другой стороны, зачастую именно сокрытие ис-

тинных целей обследования позволяет выявить интересующие исследователя психологические характеристики испытуемых. Выше уже отмечалось отсутствие очевидной валидности у проективных методик, что позволяет снизить вероятность фальсификации результатов по сравнению с опросниками. Приведём ещё один пример методики с низкой очевидной валидностью, а именно методику первичного выявления лидера группы [109].

Нескольким лицам предлагается в максимально короткий срок рассортировать по цвету и по форме картонные фишки (треугольники, круги и квадраты, окрашенные в три цвета). Цель, объявляемая группе: идет проверка на быстроту реакций и внимательность. В действительности исследователя интересует, кто из испытуемых примет на себя оперативное руководство этой совместной работой. Вначале каждый член группы пытается тасовать фишки сам за себя, потом кто-то предлагает разделить функции (одни тасуют по цвету, другие – по форме). Этот человек и проходит испытание на лидерство.

**Валидность и надёжность** – связанные понятия, различаемые с помощью понятий объекта и предмета измерения. Надёжность – это устойчивость ПД методики относительно объектов измерения (испытуемых). Надёжность не обязательно предполагает валидность. В психологии встречаются ситуации, когда вначале предлагается процедура измерения, доказывается её надёжность, как способность устойчиво различать объекты, а вопрос о валидности остаётся открытым. Валидность некоторых методик раскрывается только в результате накопления значительного опыта работы с ними. Таким образом, надёжность является необходимым, но не достаточным условием валидности и основное соотношение психометрики имеет вид:

$$\text{валидность} \leq \text{надёжность}$$

То есть валидность теста не может превышать его надёжность. Суть данного утверждения не в арифметическом соотношении, а в том, что **надёжность является необходимым, но не достаточным**

**условием валидности** (ненадёжный тест не может быть валидным, и, наоборот, валидный тест всегда надёжен).

Рассмотрим иллюстрацию связи валидности и надёжности [44]. Допустим, имеются два стрелка А и В. Стрелок А выбивает 90 очков из 100, а стрелок В – 70 очков из 100. Надёжность стрелка А – 0,9, а стрелка В – 0,7. Однако, стрелок А всегда стреляет по чужим мишеням, поэтому на соревнованиях его результаты не засчитываются. Стрелок В всегда правильно выбирает мишени. Поэтому «валидность» стрелка А нулевая, а стрелка В – 0,7, то есть равна надёжности. Если стрелок А начнёт правильно выбирать мишени, его валидность тоже будет равна его надёжности. Если же он будет продолжать иногда путать мишени, то часть результатов не будет зачтена и валидность стрелка А будет ниже его надёжности. В этом примере аналогом надёжности является меткость стрелка, а аналогом валидности – точность стрельбы по строго определённой «своей» мишени. Поэтому надёжный стрелок может стать валидным не только, сменив мишень, но и сменив команду. В истории психодиагностики известны случаи, когда тест, признанный инвалидным для измерения одних свойств, признавался валидным по отношению к другим.

### **Контрольные вопросы и задания**

1. *Рассчитайте Z-показатели для точек  $\bar{x} \pm i \cdot \sigma; i = 1, 2, 3$ .*
2. *Дайте определение надёжности психодиагностической методики и перечислите основные типы надёжности.*
3. *Охарактеризуйте основные пути установления валидности строгоформализованной методики.*
4. *Сформулируйте основное соотношение психометрики.*
5. *Определите, какой тип валидности характеризуется каждым из нижеследующих высказываний:*
  - *отражает адекватность модели тестируемой деятельности, правильность подбора заданий, охватывающих главные аспекты изучаемого феномена;*
  - *определяется путём сопоставления тестовых оценок с проявлениями измеряемого свойства в жизни;*

- оценивается по корреляции результатов теста с результатами другого теста, измеряющего то же качество, валидность которого установлена;
- отражает соответствие результатов тестирования показателям тех психологических качеств, которые теоретически связаны с оцениваемым свойством.

### 3.3. Этапы конструирования строгоформализованных методик

Конструирование строгоформализованных методик (рис. 3.9) начинается с этапа разработки методологической концепции будущей методики. В рамках концепции разрабатывается направленность методики, определяется диагностический подход и форма заданий. Всё это определяет **модель** будущей методики, включающую в себя:

- назначение методики (измеряемый конструкт) и характеристики (половозрастные, образовательные, демографические и т.д.) целевой аудитории, для которой предназначена методика;
- предположительный перечень психологических особенностей лиц, которые будут получать высокие и низкие баллы по отдельным субтестам и тесту в целом;
- общую структуру теста, (количество субтестов, шкал);
- способ конструирования общей шкалы;
- типы и формы заданий;
- количество заданий в субтестах;
- общие подходы к процедуре обследования;
- технологический подход к первичной обработке результатов;
- тип шкалы стандартных показателей.

#### Шкалирование

Способ организации эмпирических данных и упорядочивания их в числовую систему называется **шкалированием**. В более общем смысле шкалирование – это метод моделирования явлений с помощью числовых систем.



**Рис. 3.9. Этапы конструирования строгоформализованной методики**

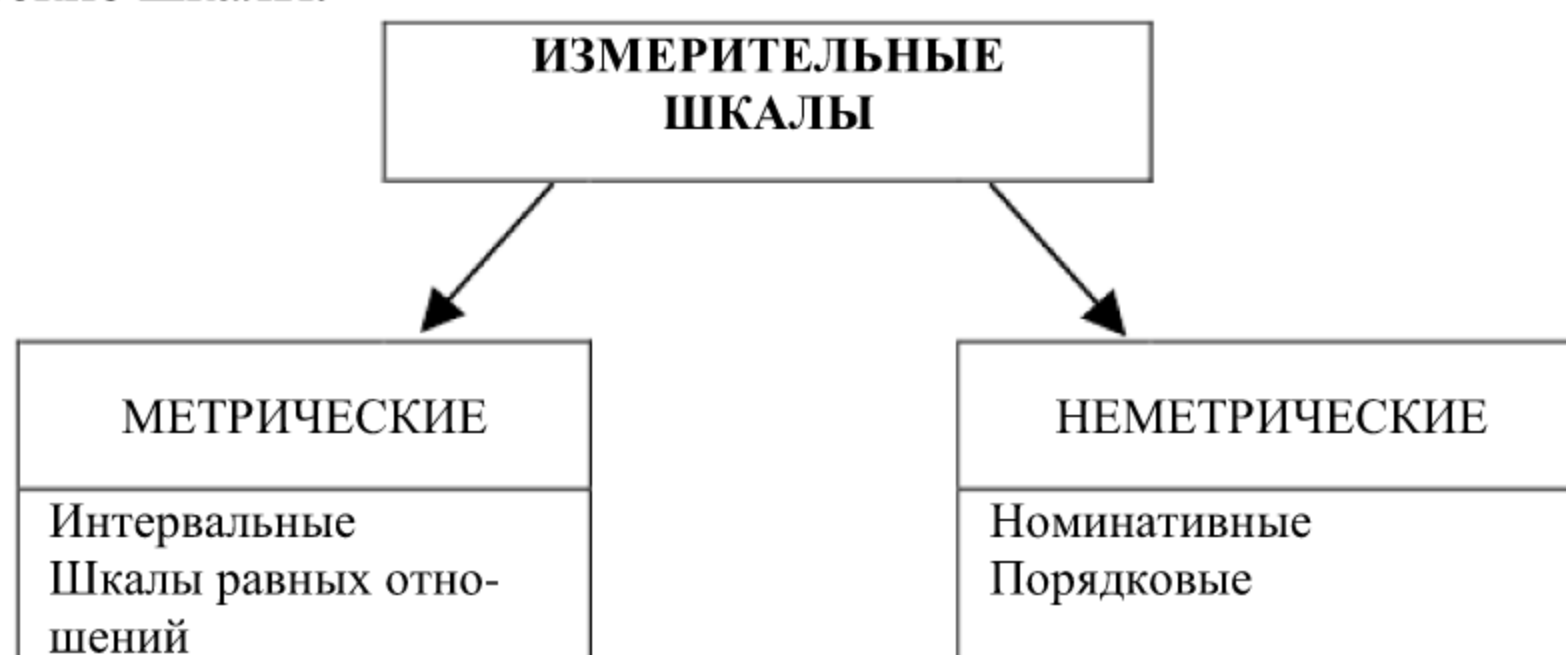
Рассмотрим шкалы измерения, используемые **при конструировании заданий** методик измерения и оценки. В главе 1 были рассмотрены следующие шкалы измерения:

- номинальная;
- порядковая;
- интервальная;
- шкала равных отношений.

Существуют следующие типы классификации шкал:

А. По признаку **формы фиксации эмпирических данных**: вербальные, числовые и графические.

Б. По признаку **метрической детерминированности** (рис. 3.10): метрические (интервальные и шкалы отношений) и неметрические (порядковые и номинативные). При **конструировании заданий** психодиагностических методик используются неметрические шкалы.



**Рис. 3.10 Классификация измерительных шкал по признаку метрической детерминированности**

При использовании номинативной шкалы испытуемый осуществляет выбор из нескольких альтернатив (такой тип заданий называется «закрытым» в отличие от «открытого» типа, при котором испытуемый сам формулирует ответ). Приведём примеры заданий из теста структуры интеллекта IST (*Intellectual Structure Test*) Р.Амтхауэра и опросника HSPQ (*High School Personal Questionnaire*) Р.Кеттелла.



**Пример 1.** Образец задания IST (субтест 1):

*Кролик больше всего похож на...*

*а) кошку; б) белку; в) зайца; г) лису; д) ежа.*

**Пример 2.** Образцы заданий HSPQ:

*1. Слово «сильный» означает то же, что и слово:*

*а) могущественный б) суровый в) выносливый*

*2. Вам хотелось бы играть на сцене, например, в школьной самодеятельности?*

*а) да б) не уверен в) нет.*

В последнем примере варианты ответов могут быть упорядочены по степени желанности испытуемого играть на сцене, однако это всё же порядковая шкала, поскольку балл в данном случае начисляется лишь при совпадении ответа с ключевым пунктом. Использование порядковой шкалы предполагает упорядочивание степени согласия испытуемого с утверждением. Упорядочивание может быть униполярным (Пример 3) и биполярным (Пример 4).

**Пример 3.** Инструкция и образец заданий опросника мотивации учебной деятельности (А.А.Реан, В.А.Якунин):

**Инструкция.**

*Оцените по 7-балльной шкале приведенные в списке мотивы учебной деятельности по их значимости для Вас и обведите в кружок соответствующее число баллов. При этом считается, что 1 балл соответствует минимальной значимости мотива, а 7 баллов – максимальной. Оценивайте все приведенные в списке мотивы, не пропуская ни одного.*

1. Стать высококвалифицированным специалистом.

1	2	3	4	5	6	7
---	---	---	---	---	---	---

2. Получить диплом.

1	2	3	4	5	6	7
---	---	---	---	---	---	---

3. Успешно продолжить обучение на последующих курсах.

1	2	3	4	5	6	7
---	---	---	---	---	---	---

4. Успешно учиться, сдавать экзамены на «хорошо» и «отлично».

1	2	3	4	5	6	7
---	---	---	---	---	---	---

5. Постоянно получать стипендию.

1	2	3	4	5	6	7
---	---	---	---	---	---	---

*и так далее.*

**Пример 4.** Инструкция и образец заданий методики СЖО (тест смысложизненных ориентаций, автор – Д.А.Леонтьев):

**Инструкция.**

*Вам будут предложены пары противоположных утверждений. Ваша задача - выбрать одно из двух утверждений, которое, по Вашему мнению, больше соответствует действительности, и отметить одну из цифр 1, 2, 3, в зависимости от того, насколько Вы уверены в выборе (или 0, если оба утверждения, на Ваш взгляд одинаково верны).*

1. Обычно мне очень скучно.	<b>3 2 1 0 1 2 3</b>	Обычно я полон энергии.
2. Жизнь кажется мне всегда волнующей и захватывающей.	<b>3 2 1 0 1 2 3</b>	Жизнь кажется мне совершенно спокойной и рутинной.
3. В жизни я не имею определенных целей и намерений.	<b>3 2 1 0 1 2 3</b>	В жизни я имею очень ясные цели и намерения.
4. Моя жизнь представляется мне крайне бессмысленной и бесцельной.	<b>3 2 1 0 1 2 3</b>	Моя жизнь представляется мне вполне осмысленной и целеустремленной.
5. Каждый день кажется мне всегда новым и непохожим на другие.	<b>3 2 1 0 1 2 3</b>	Каждый день кажется мне совершенно похожим на все другие

*и так далее.*

Интервальная шкала в психодиагностических измерениях используется при выражении результатов в стандартных показателях (Т-баллы, IQ – шкала и т.п., см. параграф 3.2). **При конструировании заданий строгоформализованных методик интервальные шкалы не используются.**

Важнейшими характеристиками качества шкалы измерения в психодиагностике считаются одномерность, надёжность и обоснованность (валидность) шкалы. Под **одномерностью** понимается сопоставимость шкалируемых параметров, отсутствие их смещений (пропорциональность между положительным и отрицательным полюсами шкалы), равенство интервалов шкалы, симметричность различных позиций. **Надёжность** шкалы определяется на основе анализа устойчивости повторных измерений. **Обоснованность** (ва-

лидность) шкалы – приспособленность данной шкалы для измерения данного психологического качества, о полноте его отражения и техническом соответствии самой процедуры шкалирования.

В основе разработки заданий для любой методики лежит теоретическая концепция, принятая автором методики в качестве методологической основы теста. В рамках этой теории определяются различные стороны изучаемого психологического качества. Задания разрабатываются «по блокам», соответствующим этим компонентам. По мнению П.Клайна [29, с.50] при разработке тестов первоначально следует сформулировать как можно большее количество заданий.

**Генеральная совокупность заданий теста** – гипотетическое множество всевозможных тестовых заданий, связанных общим фактором (измеряющих заданный психологический конструкт). Каждый тест теоретически представляет собой выборку из генеральной совокупности заданий, которая релевантна исследуемой психологической переменной (так, тест тревожности должен представлять выборку из всех возможных заданий, с помощью которой может быть обнаружена эта особенность личности). Если отобранные для теста задания не репрезентативны генеральной совокупности заданий, то тестовые результаты будут искажены. Анализ этой репрезентативности является компонентом конструктивной валидности методики. Далее будут рассмотрены правила разработки заданий для следующих категорий строгоформализованных методик: когнитивные тесты (тесты общих и специальных способностей, тесты достижений) и личностные опросники.

### **Задания для тестов интеллекта**

Рассмотрим примеры заданий, наиболее часто использующихся в тестах интеллекта. Отметим, что при разработке теста в первую очередь нужно опираться на модель интеллекта, выбранную в качестве методологической основы теста. Во-вторых, содержание заданий должно соответствовать той категории испытуемых, для которых предназначен тест.

В тестах интеллекта наиболее часто используются задания следующих типов [29, с.52: аналогии, группировки («исключи лишнее»), последовательности. Приведём примеры этих заданий из теста структуры интеллекта Р.Амтхауэра.

## **1. Аналогии.**

### ***Инструкция.***

*В каждом задании Вам предлагается две пары слов, при этом во второй паре одного слова не хватает. Первая пара слов – полная, состоящая из двух взаимосвязанных по смыслу слов; нужно понять смысл этой взаимосвязи, чтобы в соответствии с ним выбрать недостающее во второй паре слово из пяти слов, приводимых ниже.*

**Образец 1.** Лес – дерево; луг – ...?

а) куст; б) пастбище; в) трава; г) сено; д) тропинка.

Ответ 1 в, т. к. взаимное отношение леса и деревьев имеет такой же смысл, как взаимное отношение луга и травы.

**Образец 2.** Темный – светлый; мокрый – ...?

а) дождливый; б) сырой; в) пасмурный; г) влажный; д) сухой.

Ответ 2д, т. к. взаимное отношение темного и светлого имеет такой же противоречивый смысл, как взаимное отношение мокрого и сухого.

## **2. Классификации.**

### ***Инструкция.***

*Вам предлагаются наборы 5 слов, 4 из которых могут быть объединены в одну группу по общему смыслу, подходящему для всех этих четырех слов. Пятое, лишнее по смыслу слово и должно быть вашим ответом на задание, которое иначе может быть названо так: «Найдите лишнее слово, не подходящее по смыслу к четырем другим из пяти названных». Это лишнее слово обозначено соответствующей буквой, которую необходимо проставить рядом с номером задания.*

**Образец 1.** а) стол; б) стул; в) голубь; г) диван; д) шкаф.

Ответ 1в, т. к. «голубь» не относится к предметам мебели, а смысл объединения слов именно таков. Слово «голубь» является лишним по смыслу в ряду названных слов.

**Образец 2.** а) идти; б) мчаться; в) ползти; г) бежать; д) лежать.

Ответ 2д, т. к. «лежать» не относится к обозначению способов передвижения. Слово «лежать» является лишним по смыслу в ряду названных слов.

### **3. Последовательности.**

#### ***Инструкция.***

*В этом разделе каждое задание представлено рядом чисел, расположенных в определенной связи между собою. Необходимо продолжить числовой ряд на основании обнаруженной вами особенности этой связи чисел.*

**Образец 1.** 2, 4, 6, 8, 10, 12, 14...

Ответ: 16, т. к. в этом ряду чисел особенность их связи между собою – постоянное возрастание каждого следующего числа на 2 единицы.

**Образец 2.** 2, 9, 7, 10, 8, 11, 9, 12...

Ответ: 10, т. к. в этом числовом ряду особенность связи чисел между собою состоит в том, что при переходе от первого числа ко второму нужно вычесть 2 единицы, а при переходе от второго числа к третьему нужно прибавить 3 единицы и т. д.

Рассмотрим далее общие правила конструирования заданий для тестов интеллекта на примере заданий типа «Аналогии»:

1. Определить отношение, с которым будет связано очередное задание.

2. Выбрать материал для задания, который должен быть настолько простым, насколько это возможно. Трудность задания должна содержаться исключительно в природе отношений, которые необходимо раскрыть, а не в узкой специфичности использованного материала.

3. Разработать задание. Существуют различные типы заданий для тестов интеллекта: задания с несколькими вариантами выбора (т.н. закрытого типа) – в большинстве случаев наиболее эффективные, альтернативные задания (да/нет), задания на восстановление соответствия – эффективные при очень подробной проверке знаний. Если при формулировке задания применяется открытая форма ответа, зачастую результат выполнения такого задания трудно оценить однозначно (испытуемый может дать «наполовину правиль-

ный ответ»). Выбор наиболее эффективного типа задания определяется его содержанием.

4. Убедиться, что компоненты задания не образуют других отношений, приводящих к альтернативным решениям.

При конструировании заданий с несколькими вариантами выбора особое внимание следует обращать на формулировку **дистракторов** (варианты ответов, не являющиеся верными). **Дистракторы для тестов интеллекта** должны относиться к той же категории, что и правильный ответ, но не быть слишком близкими к правильному ответу (практически неотличимыми от него). **Дистракторы для тестов достижений** должны быть такими, чтобы каждый из них мог привлечь внимание испытуемых (то есть быть как бы «похожими» на правильный ответ). Так, в ЕГЭ дистракторы формулируются с учётом наиболее типичных ошибок учеников.

### **Задания для личностных опросников**

Как и в случае тестов способностей и достижений, качество отдельных заданий (вопросов, утверждений) является определяющим для всего опросника. При разработке заданий для личностных опросников необходимо учитывать следующие проблемы, которые, если их не обойти, могут привести к низкой валидности опросника [29].

1. Установка на согласие. Проявляется, когда вопросы неоднозначны и неопределённые.

2. Установка на социально одобряемые ответы.

3. Установка на неопределённые или средние ответы, приводящая к снижению валидности заданий, поскольку методы анализа заданий основываются на крайних значениях показателей.

4. Очевидная валидность вопросов (заданий).

5. Трудность определения конкретной выборочной совокупности вопросов, связанная с проблемой отнесения вопроса к генеральной совокупности вопросов, направленных на диагностику определённого личностного качества (в отличие от тестов интеллекта и способностей, где принадлежность вопросов к той или иной ге-



неральной совокупности определить легко: лингвистические задания трудно перепутать с арифметическими). Кеттелл отмечает, что вопросы часто бывают нагружены не тем фактором, для выявления которого они предназначались, либо вообще не нагружены ни одним фактором. Указанная проблема подбора адекватной выборки из выборочной совокупности вопросов или утверждений для опросника напрямую связана с его валидностью.

6. Проблема в установлении адекватного критерия валидации.

### **Правила формулировки заданий для опросников**

1. Перечислить черты и особенности поведения, которые предполагается изучать с помощью опросника.

2. Преобразовать описание каждой черты или особенности поведения в задание (множество заданий). Выбор формы каждого задания зависит во-первых, от особенностей материала и во-вторых от преимуществ и недостатков каждого типа заданий:

- 1) дихотомические (да/нет)
- 2) трихотомические (да-?-нет)
- 3) альтернативные
- 4) «нравится – не нравится»
- 5) вынужденный выбор
- 6) задание с рейтинговой шкалой

3. Формулировка каждого задания должна быть такой, чтобы минимизировать проникновение испытуемых в сущность того, что изучается при помощи этого задания.

4. Формулировка заданий должна быть простой и недвусмысленной.

5. Формулировка заданий должна быть тщательно продуманной для минимизации установки на социально одобряемые ответы.

6. Шкала должна быть сбалансированной (равенство числа ответов «да» и «нет» в ключе). Это уменьшает искажение результатов вследствие установок на согласие.

7. Каждое задание должно относиться к только одной области поведения.

8. В каждом задании должен содержаться только один вопрос или утверждение.

9. Рекомендуется не использовать термины частоты и другие субъективные выражения.

10. По возможности задания должны касаться поведения, а не чувств и отношений.

### **Отбор и оценка заданий**

После того, как разработано достаточное множество заданий (в несколько раз больше, чем необходимо для создаваемой методики в соответствии с моделью), производится отбор и оценка заданий для конструирования шкалы. Отбор производится в два этапа:

1. На этапе первоначального конструирования шкалы производится первоначальный («умозрительный») отбор и оценка заданий разработчиками методики.

2. После проведения пилотажного тестирования проводится оценка дискриминативности каждого задания, т.е. соответствие результата его выполнения общему результату методики (см. параграф 3.2), а также производится пересмотр формулировки тех заданий, которые в ходе пилотажного тестирования вызвали вопросы и замечания испытуемых и экспериментаторов.

Как правило, в ходе апробации методики вышеперечисленные этапы повторяются не один раз.

### **Разработка инструкции**

Инструкция к ПД методике должна быть (П.Клайн):

- краткой, насколько это возможно;
- включать только простые предложения без уточняющих придаточных предложений;
- максимально понятной, с учётом особенностей выборки;
- примеры должны помогать понять инструкцию.

П.Клайн рекомендует знакомить испытуемых со всеми инструкциями до тестирования, а не во время его, вплоть до предъявления

им простейших заданий, дающих возможность удостовериться в том, что все инструкции поняты. В первую очередь, эта рекомендация относится к когнитивным тестам (способностей и достижений). Тогда во время тестирования, во-первых, у экспериментатора будет уверенность в том, что испытуемые поняли инструкцию, а во-вторых, будет установлен хороший контакт с испытуемыми, что существенно, так как при тестировании интеллектуальных способностей от испытуемых требуется продемонстрировать максимум возможностей.

### **Пилотажное тестирование**

После того, как отобраны задания, сконструирована шкала и разработана инструкция, производится оформление буклетов заданий, бланков и других составляющих инструментария. Методика готова к первому пилотажному апробированию. На этом этапе полезно также разработать сценарий всей процедуры пилотажного обследования.

Выборка для проведения пилотажного тестирования не должна быть очень большой. Вполне достаточна группа объёмом в 20 человек, но состав группы должен соответствовать тому контингенту испытуемых, для которых предназначен тест. Экспериментатор проводит обследование выборки в соответствии с разработанным сценарием, соблюдая все правила проведения психодиагностического обследования. Ассистент экспериментатора записывает все вопросы, которые возникали у испытуемых в ходе обследования, а также все замечания по процедуре.

По результатам пилотажного тестирования оцениваются коэффициенты дискриминативности заданий, производится анализ и устранение всех замечаний, возникших в ходе обследования (как по процедуре проведения, так и по содержанию и оформлению заданий). На этом этапе также отрабатывается технология обработки результатов и разрабатываются инструкции для психологов-пользователей методики.

В случае необходимости пилотажное обследование повторяется.

Итогом пилотажной апробации является инструментарий, полностью подготовленный к процедуре сбора нормативных данных и разработки шкалы стандартных показателей.

Отметим, что все вышеперечисленные этапы разработки строгоформализованных методик применяются и при разработке анкет.

### **Стандартизация оценки выполнения теста**

Содержание процедуры стандартизации оценки выполнения теста включает сбор нормативных показателей, конструирование шкалы и составление формулы (таблицы) перевода сырых баллов в стандартные показатели. Содержание этих этапов подробно рассмотрено в параграфе 3.2, поэтому здесь ограничимся описанием общей схемы действий.

Первоначально формируется выборка стандартизации, подбор которой следует осуществлять с большой тщательностью. В соответствии с моделью определяются основания стратификации, и затем из каждой страты (группы) методом случайного отбора подбираются участники выборки стандартизации.

Далее наступает организационный период процедуры стандартизации: в нужном количестве издаются буклеты и бланки для испытуемых, а также методические рекомендации для специалистов, которые будут проводить обследование. Желательно, чтобы эти специалисты первоначально сами прошли обследование по данной методике, а после этого отработали навыки проведения обследования на небольшой (от 3-5 чел.) группе испытуемых, не входящих в выборку стандартизации. В этом случае резко снижается вероятность недостоверности результатов обследования выборки стандартизации вследствие нарушений регламентации.

После получения результатов обследования выборки стандартизации, их обработки и оцифровки производится проверка на нормальность. Технология проверки эмпирических данных описана в главе 6.

В том случае, когда распределение полученных результатов обследования выборки стандартизации статистически значимо отли-

чается от нормального, необходимо разобраться в характере отклонений. Если, например, речь идёт о тесте интеллекта и характер отклонений от нормального распределения представляет собой левостороннюю асимметрию (рис. 3.11), то, очевидно, что задания теста оказались слишком трудными для значительной части обследованной выборки. Если асимметрия правосторонняя (рис. 3.12), то ситуация обратная: в тесте оказалось много слишком лёгких заданий. В подобных случаях требуется рассчитать индекс трудности заданий и пересмотреть (сократить долю в общем числе заданий) те из них, которые имеют слишком низкий (слишком высокий) индекс трудности. После этого (рис. 3.9) требуется проведение новой процедуры стандартизации.



**Рис. 3.11 Левосторонняя асимметрия**



**Рис. 3.12 Правосторонняя асимметрия**

Если по ряду причин не удаётся добиться нормальности распределения признака в выборке стандартизации, то требуется либо пересмотр первоначальной концепции и модели, либо (например, в том случае, когда проводится рестандартизация методики и содержание заданий изменить невозможно) использовать процентиля.

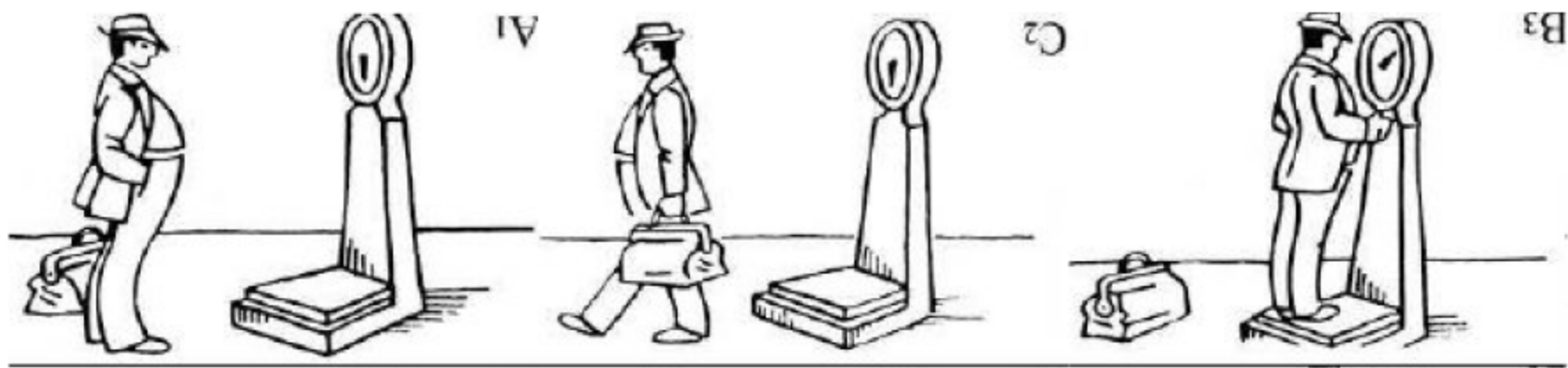
Перевод сырых баллов в шкалу стандартных показателей осуществляется только в том случае, когда данные обследования выборки стандартизации имеют нормальный закон распределения. Технология перевода подробно описана в параграфе 3.2.

### **Психометрическая проверка (надёжность, валидность)**

Следующим этапом разработки профессиональной строгоформализованной методики является психометрическая проверка. В соответствии с теоретической концепцией и моделью методики подбираются наиболее подходящие типы валидности и виды надёжности (см. параграф 3.2).

Вопрос выбора относится к компетенции разработчика. Многие из существующих типов валидности пригодны исключительно для определённого типа методик. По сути, как уже отмечалось выше, доказательство валидности, т.е. пригодности теста можно проводить различными путями. Соответственно получаются и разные типы валидности. Так, например, неким эталоном для тестов интеллекта в течение достаточно продолжительного времени являлся тест Д.Векслера. Авторы многих вновь разрабатываемых методик считали своим долгом указать данные о конкурентной валидности своих тестов со шкалами теста Векслера. Вместе с тем, несмотря на то, что процедура исследования конкурентной валидности достаточно проста, необходимо обращать внимание на 1) методологические основы сопоставляемых методик; 2) влияние исторической эпохи на содержание стимульного материала теста: так, например, неверное выполнение задания (упорядочивание картинок) детского теста Векслера (рис. 3.13) может быть связано с непониманием испытуемого назначения устройства (уличных весов), не используемого в настоящее время.





**Рис. 3.13 Стимульный материал одного из субтестов теста Векслера**

После проведения психометрической проверки, получения данных о валидности и надёжности методики можно публиковать разработанную методику. Методическое руководство должно включать следующую информацию:

- методологическая концепцию, лежащую в основе методики;
- структура методики, назначение и психологическое содержание шкал;
- все данные о психометрической проверке (объём выборок, тип показателя, процедура эксперимента, результаты);
- подробную инструкцию для пользователей теста по проведению и первичной обработке результатов обследования;
- данные о выборке стандартизации (выборки стандартизации может быть несколько в тех случаях, когда стандартизация проводилась на выборках, которые не могут быть объединены);
- тип стандартных показателей и таблицу перевода сырых показателей в стандартную шкалу;
- этапы и примеры построения психологического заключения, характеристику испытуемых, набравших высокие (низкие) баллы по каждой шкале, а также сочетание высоких и/или низких баллов по шкалам;
- рекомендации по использованию методики в различных областях психологической практики.

С течением времени необходимо проводить рестандартизацию и пересмотр заданий методики.

### **Контрольные вопросы и задания**

1. *Перечислите этапы конструирования строгоформализованной методики и кратко раскройте их основное содержание.*
2. *Какие типы шкал используются при конструировании строгоформализованных методик. Приведите примеры, отличные от примеров в тексте параграфа.*
3. *Приведите примеры типов заданий тестов интеллекта, не рассмотренные в настоящем параграфе.*
4. *Перечислите проблемы разработки заданий для личностных опросников.*
5. *Какой должна быть инструкция к строгоформализованной методике?*
6. *С какой целью проводится пилотажное тестирование?*
7. *Приведите примеры устаревших, по вашему мнению, заданий из известных психодиагностических методик. Обоснуйте своё мнение.*
8. *Какая информация о строгоформализованной методике должна присутствовать в методическом руководстве?*

### **3.4 Адаптация зарубежных строгоформализованных методик**

В 90-е годы XX века в нашей стране произошёл «тестовый бум»: массовыми тиражами начали издаваться переводные западные тесты, опросники, проективные методики. Эти тесты объединялись в сборники под названиями, обеспечивающими коммерческий успех изданию: «Тесты для женщин», «Альманах лучших тестов» и т.п. Каждый тест из такого сборника представлял собой набор заданий, ключи для обработки и краткую характеристику результата в зависимости от количества набранных баллов. Никакой другой информации не предоставлялось.

Зачастую эти методики переводили на русский язык переводчики, не являющиеся профессионалами в области с психодиагностики и психометрики. Смысл некоторых фраз при таком переводе существенно искажался. Помимо этого, не принималось во внимание то, что перевод теста, каким бы профессиональным он ни был, – только начало достаточно длительного и трудоемкого процесса адапта-

ции методики. Понадобилось без малого десять лет, для того чтобы сформировалась элементарная психометрическая культура пользователей тестов. Тем не менее, как отмечает Л.Ф.Бурлачук [10], еще рано говорить о том, что в деле адаптации психологических тестов у нас наведен порядок.

Редко какой зарубежный тест, используемый в СНГ, по словам Л.Ф.Бурлачука, соответствует всем этим требованиям профессиональной методики. Преобладающее большинство из применяемых у нас зарубежных тестов не может быть рекомендовано многочисленным психологам-пользователям, а значение получаемых с их помощью данных, как правило, не выходит за пределы тех конкретных научных исследований, в которых они были использованы.

Наименее сложны в адаптации проективные методики, стимульный материал которых не подлежат каким-либо изменениям. Это же можно сказать и о многих других невербальных тестах (но не о всех: так, в широко распространённой в России методики экспресс-диагностики интеллектуальных способностей МЭДИС в стимульном материале встречаются изображения бейсбольной биты, мотыги и т.п., типичные для американской аудитории, но вызывающие вопросы у российских испытуемых). Наибольшие сложности вызывает адаптация широко распространенных в отечественной психодиагностике личностных опросников.

Процесс их адаптации нередко ограничивают следующими этапами [10]:

- подготовка предварительного варианта перевода с оригинала на второй язык;
- экспертная содержательная оценка предварительного перевода с привлечением лингвистов и профессиональных психологов, владеющих языком оригинала;
- проверка эквивалентности экспериментального перевода шкалы и оригинала;
- установление новой шкалы и сбор соответствующих норм на отечественной выборке.

Перевод с языка оригинала (чаще всего – английского) на русский язык сопряжён с определёнными семантическими трудностями. Д. Кэмпбелл, различая два вида перевода – симметричный и асимметричный, отмечает обычное для перевода опросников желание сохранить верность оригиналу, а в итоге «на свет появляется неуклюжая, экзотическая версия теста». Это происходит потому, что «образный, метафорический перевод, при котором идиомы языка, на котором осуществляется перевод, заменяют собой аналогичные идиомы оригинала, не допускался из боязни нарушить взаимооднозначную идентичность вопросов» (Цит. по [10]). Каждому специалисту, работающему в области психологической диагностики, неоднократно приходилось сталкиваться с такими «экзотическими» опросниками, с помощью которых, как точно отмечает Л.Ф.Бурлачук, могут быть получены данные скорее о степени понимания испытуемым сложных грамматических конструкций, нежели об особенностях его личности.

Лингвистический аспект адаптации личностных опросников означает приспособление лексики и грамматики к возрастной и образовательной структуре населения, для которого эта методика предназначена, учет коннотативного значения языковых единиц и категорий.

Приведём пример двух различных переводов одного и того же вопроса из опросника Х.Смишека:

Вариант 1. Часто ли Вы, стоя у окна многоэтажного жилого дома испытываете желание выпрыгнуть из него?

Вариант 2. Бывает ли, что Вы, стоя у окна многоэтажного дома, остерегаетесь того, что можете выпасть из окна?

Данный вопрос относится к педантичному типу акцентуации, и, безусловно, первый вариант перевода является абсолютно неподходящим.

Л.Ф.Бурлачук также подчёркивает, что перевод любого опросника – это не только сложная лингвистическая задача. В личностных опросниках всегда отражены особенности культуры того общества, в котором они создавались. Нормам и ценностям одного

общества трудно, а порой и невозможно найти эквиваленты в другой культуре.

Следующий этап адаптации – психометрическая проверка, которая обязательно должна включать следующие исследования, относящиеся к валидности и надёжности методики:

1) анализ внутренней согласованности вопросов (утверждений), из которых составлена шкала, что обязательно по отношению к факторным опросникам;

2) проверка ретестовой надёжности;

3) оценка критериальной валидности (внешний критерий должен быть релевантным измеряемому конструкту).

4) для многофакторных опросников – анализ воспроизводимости структуры взаимосвязи между факторами.

После психометрической проверки проводится стандартизация по изложенной выше схеме.

В заключение параграфа приведём слова Л.Ф.Бурлачука [10], что профессиональная адаптация многих личностных опросников по своему объёму не уступает разработке оригинальной методики. Поэтому ещё до начала адаптации зарубежного теста необходимо быть уверенным в том, что именно данный тест необходим и восполняет определенные пробелы в диагностическом инструментарии. Поэтому адаптации теста должно предшествовать изучение его теоретических оснований. В идеальном варианте на смену адаптации зарубежных тестов должна прийти разработка оригинальных отечественных методик.

### **Контрольные вопросы и задания**

1. *Перечислите этапы адаптации зарубежной методики и раскройте их содержание.*
2. *Приведите примеры симметричного и асимметричного перевода зарубежных методик.*

## Глава 4.

### Малоформализованные методы психологической оценки

#### 4.1. Метод наблюдения

Метод наблюдения является старейшим методом психологической диагностики. С его помощью можно получить обширную информацию о человеке. Особо важное значение метод наблюдения имеет для изучения психологических особенностей детей, поскольку ребёнок как объект исследования представляет бóльшие трудности для экспериментального изучения, чем взрослый человек.

Метод наблюдений в детской психологии и педагогике был научно разработан в нашей стране в 20-е годы Михаилом Яковлевичем Басовым [73]. Научное наблюдение как психодиагностический метод характеризуется:

- постановкой проблемы,
- выбором ситуаций для наблюдения,
- определением психологических качеств или особенностей поведения, которые должны стать объектом наблюдения,
- разработанной системой фиксации и записи результатов.

Наблюдение бывает поисковым и исследующим. Поисковое наблюдение производится обычно на начальном этапе разработки какой-либо проблемы. При этом наблюдается всё, чем характеризуется объект, без отбора каких-либо определённых его проявлений.

Если цель наблюдения конкретна и определена, то в этом случае производится отбор только нужных фактов и явлений. Такое наблюдение называется исследующим или выбирающим. Здесь заранее определено предметное содержание наблюдения (что наблюдать) и проведено расчленение наблюдаемого на единицы.



Независимо от того, какой характер носит наблюдение – поисковый или исследующий, наблюдатель должен иметь определённую схему действий. Схема наблюдения включает перечень единиц наблюдения, способ и форму описания наблюдаемого явления.

Определение единиц поведения при поисковом наблюдении сложнее, чем при исследовательском. Так, наблюдая за поведением ребёнка вообще, его следует разделить на ряд единиц: моторика, речь, общение, эмоции и т.д. Если же предметом наблюдения является только речь ребёнка, то единицами могут быть содержание речи, её направленность, особенности лексического, грамматического, фонетического строя и т.п.

Выбор способов и формы описания наблюдения зависят от того, каков его характер: поисковый или исследующий. Вместе с тем, имеются общие требования к записи наблюдения [73].

1. Запись должна фиксировать наблюдаемый факт в том виде, в котором он существовал реально, не подменяя его описанием личных впечатлений и разнообразных суждений самого наблюдателя. То есть записывать нужно только что происходило и каким образом (фотографическая запись).

2. Запись должна фиксировать не только наблюдаемый факт, но и ту окружающую обстановку (фон), в которой он происходил.

3. Запись должна по возможности полно и в соответствии с целью отражать изучаемую реальность.

При поисковом наблюдении используются формы записей в виде сплошного протокола или дневника. Вести запись желательно во время наблюдения. Если это не всегда возможно, то следует зафиксировать существенные моменты, а подробности – сразу же после окончания наблюдений. Обработку результатов нужно производить сразу после окончания наблюдения, в тот же день.

При исследующем наблюдении протокол чаще всего имеет вид заранее разработанной таблицы, в который вносятся не содержательное описание единиц поведения, а символы (буквы, цифры), которые обозначают определённую категорию. Для уменьшения субъективизма наблюдателя при описании и обработке результа-

тов, используется психологическое шкалирование, направленное на оценку степени выраженности наблюдаемых признаков.

### **Виды наблюдения.**

#### **1. Наблюдения хронологические:**

- лонгитюдные (в течение нескольких лет, предполагают постоянный контакт исследователя и объекта изучения);
- периодические (проводятся в течение определённых промежутков времени),
- единичные (описание отдельного случая).

Пример В.С.Юркевич: наблюдение через зеркало Гезелла детско-родительских отношений (в игровой комнате) с последующим сопоставлением с результатами диагностики одарённости ребёнка.

#### **2. В зависимости от ситуации:**

- полевые (естественными для объекта наблюдений);
- лабораторные (объект наблюдается в искусственных условиях);
- спровоцированные в естественных условиях.

#### **3. В зависимости от позиции наблюдателя:**

открытое или скрытое (через зеркало Гезелла);  
наблюдение со стороны или включённое (исследователь является членом группы, полноправным её участником).

Включённое наблюдение может быть открытым и скрытым (наблюдатель действует инкогнито).

Правила эффективности метода наблюдений [75].

- проводить многократные систематические наблюдения данного факта в разнообразных ситуациях, что позволит отделить случайные совпадения от устойчивых закономерных связей;
- не делать поспешных выводов, обязательно выдвигать и проверять альтернативные предположения относительно той реальности, которая стоит за наблюдаемым фактом;
- не отрывать частные условия возникновения наблюдаемого факта от общей ситуации, рассматривать их в контексте общей ситуации;
- стараться быть беспристрастным;

- оценивать одного обследуемого должны несколько наблюдателей (не менее 2-х), при этом их суждения должны быть независимыми.

Составление жёсткой схемы наблюдения, шкалирование оценок способствует повышению объективности и достоверности информации. Прежде, чем предлагать нескольким экспертам проводить наблюдение, необходимо чётко определить, какие поведенческие проявления стоят за каждым определением. Вместе с тем, полностью исключить влияние личности экспериментатора на результаты наблюдения невозможно.

### **Контрольные вопросы и задания**

1. *Охарактеризуйте метод научного наблюдения М.Я.Басова*
2. *Каковы формы записи протокола при поисковом и исследующем наблюдении?*
3. *Реализуйте программу наблюдения за интересом студентов к занятиям. Охарактеризуйте поведенческие проявления следующих уровней проявления интереса:*
  - *отсутствует (1)*
  - *едва проявляется (2)*
  - *проявляется (3)*
  - *проявляет большой интерес (4)*
  - *проявляет жгучий интерес (5)*

## **4.2. Метод анкетного опроса**

Метод анкетного опроса применяется при изучении субъективных состояний людей, мотивации, мнений и отношения к событиям. Анкетой называется перечень стандартизированных вопросов с возможными вариантами ответов или без них, которые подаются респонденту в печатном виде.

Опросы имеют большие преимущества перед другими методами сбора данных, поскольку позволяют получить информацию от большого количества респондентов в сжатые сроки. Достоверность

полученных данных обеспечивается научной обоснованностью тщательностью выполнения основных этапов проведения анкетного опроса:

- подготовительный (определение целей и задач исследования, формулировка вопросов и составление анкеты);
- оперативный (реализация процедуры сбора информации с помощью разработанных анкет);
- результирующий этап обработки и анализа полученных данных.

Для проверки обоснованности и надёжности данных, полученных в результате анкетного опроса, а также для повышения достоверности и более глубокой интерпретации данных анкетный опрос рекомендуется комбинировать в одном исследовании с интервью (для интервью отбирается, как правило 5-6% респондентов) и с наблюдением.

Анкета заполняется респондентом самостоятельно, поэтому принцип работы с анкетой должен быть простым и понятным. Надёжность данных существенно зависит от конструкции вопросов, которые подразделяются на открытые и закрытые.

### **Типология вопросов анкеты**

*Открытые вопросы* (свободная форма ответа) позволяют выявить доминанту мнений, оценок, настроений: люди отмечают те стороны явлений или говорят о том, что волнует их больше всего, о том, что доминирует в их сознании. Но самое главное, как отмечает В.А.Ядов [109], состоит в том, что, реагируя на вопрос без подсказки вариантов ответа, люди лучше проявляют особенности своего повседневного, обыденного сознания, свой образ мыслей.

Например, при изучении особенностей образовательного процесса с помощью анкетного опроса студентов и преподавателей, в анкетах как для студентов, так и для преподавателей был следующий вопрос с открытой формой ответа: *«Перечислите по 3 положительных и отрицательных качества у преподавателя»*.

Анализ ответов преподавателей на этот вопрос показал большую вариативность ответов: число различных терминов, характеризующих положительные и отрицательные стороны преподавателей оказалось сопоставимым с объёмом выборки. Поэтому их ответы были разбиты на следующие категории:

- когнитивные характеристики;
- педагогические характеристики и ораторское мастерство;
- личностные качества и мотивация;
- внешний вид.

В результате анализа результатов выяснилось, что в числе положительных сторон преподавателя вуза 30% респондентов-преподавателей отметили когнитивные характеристики, 14% – педагогические характеристики и ораторское мастерство, 2% – внешний вид, и подавляющее большинство характеристик – 54% – относились к категории «личностные качества и мотивация». Аналогично в числе отрицательных сторон преподавателя вуза 16% респондентов-преподавателей отметили когнитивные характеристики, 7% – педагогические характеристики и ораторское мастерство, 4% – внешний вид, и подавляющее большинство характеристик – 73 % – относились к категории «личностные качества и мотивация».

Главный недостаток открытых вопросов, как отмечает В.А.Ядов, состоит в том, что «высказываемые здесь мнения и оценки связаны с какими-то неизвестными нам рамками сравнения, которые очерчивают контекст высказанных суждений. Изменение границ сопоставления суждений ведет к изменению акцентов: доминирующие пункты информации могут оказаться на периферии, периферийные – передвинутся в центр внимания опрашиваемого».

Второй недостаток открытых вопросов заключается в большей сложности и трудоёмкости обработки данных. Свободная форма ответов предполагают последующую группировку, квантификацию и контент-анализ ответов. Кроме того, при обработке свободных ответов требуется «искусство расшифровки» реального смысла, который респондент вложил в ответ – т.е. индивидуальная кропотливая работа с каждой анкетой.

*Закрытые вопросы* позволяют более строго интерпретировать ответ. Рамки соотнесения оценок и суждений определяются здесь набором единых для всех опрошенных вариантов ответа. Исследователь имеет более надежные основания, чем при открытых вопросах, сопоставлять данные в равных условиях. Кроме того, существует возможность измерить интенсивность оценок, шкалируя их по каждому варианту.

Приведём пример вопроса закрытого типа из анкетного опроса преподавателей и студентов.

*4. Выберите наиболее подходящий вариант завершения предложения. Если в вузе будут работать курсы повышения педагогического мастерства ...*

*4.1 я постараюсь их посещать.*

*4.2 вряд ли у меня найдётся достаточно времени для их посещения.*

*4.3 мне не нужно повышать педагогическое мастерство.*

В случае шкалирования каждого варианта ответа, тот же вопрос анкеты мог бы выглядеть следующим образом:

*4. Отметьте степень своего согласия с каждым вариантом завершения предложения (1 соответствует минимальной степени согласия, а 7 – максимальной).*

*Если в вузе будут работать курсы повышения педагогического мастерства ...*

4.1 я постараюсь их посещать.

1	2	3	4	5	6	7
---	---	---	---	---	---	---

4.2 вряд ли у меня найдётся достаточно времени для их посещения

1	2	3	4	5	6	7
---	---	---	---	---	---	---

4.3 мне не нужно повышать педагогическое мастерство.

1	2	3	4	5	6	7
---	---	---	---	---	---	---

Постановка закрытых вопросов предполагает соблюдение следующих основных требований [109].

1. Максимально предусмотреть возможные варианты ответов. Используют также *полузакрытый* вариант, в котором оставляется линия для дополнительных комментариев и замечаний.



### **Пример.**

**6. Какие активные методы обучения Вы используете?**

*6.1 Программированное обучение*

*6.2 Проблемное обучение (проблемные лекции, вопросы, задания – подчеркните)*

*6.3 Интерактивное обучение (эвристическая беседа, дискуссия, мозговой штурм, метод «круглого стола», деловая игра, конкурсы практических работ с их обсуждением – подчеркните)*

*6.4 Внеаудиторные формы занятий (лекции с занимательной тематикой, тематические вечера, факультативные занятия, небольшие экспериментальные исследования – подчеркните)*

*6.5 \_\_\_\_\_*

Важно отвести должное место для *комментария и уточнений*. Если в анкете не предусмотрено достаточное место для ответов на открытый вариант вопроса, он не работает.

2. Формулируя варианты ответов, следует помнить **три важных правила**, подтверждаемых экспериментальными исследованиями:

- Отвечающий на вопрос чаще выбирает первые варианты ответа, реже – последующие. Поэтому **первыми должны быть наименее вероятные варианты ответа**.
- Чем длиннее вариант ответа, тем меньше вероятность его выбора, так как для усвоения смысла требуется больше времени, а респондент не склонен его тратить. Поэтому **подсказки должны быть примерно равной длины**.
- Чем более общий (абстрактный) характер вариант ответа, тем меньше вероятность его выбора. Поэтому **все варианты ответов следует выдерживать на одном уровне конкретности** (например, спрашивая об отношениях в коллективе, уточнить: «в Вашем коллективе в данное время»).

3. Нельзя комбинировать несколько идей в одной фразе, например: «работа интересная и хорошо оплачивается». Лучше перечислить оба признака и предложить оценить их значимость по шкале интенсивности.

4. Все возможные варианты ответов должны быть отпечатаны на одной странице, чтобы респондент мог охватить их одним взглядом.

5. Нельзя печатать всю серию положительных вариантов ответа подряд и следом за нею – серию отрицательных, или наоборот. В этих случаях мнение навязывается самой последовательностью предложенных вариантов.

6. Список предложенных ответов иногда бывает столь обширным, что опрашиваемые устают по мере продвижения к его концу и с последними группами суждений работают менее внимательно, чем с первыми, или же начинает действовать сила инерции в ответах. В таком случае В.А.Ядов предлагает расчленить список на три блока и предложить части опрашиваемых блокировку в одной последовательности, остальным группам – в другой. Например, перечисляются рубрики газеты (всего 21 наименование), и опрашиваемые должны ранжировать их в шкалах от «постоянно читаю» до «не читаю вообще». Разобьем список из 21 наименования на три части: (а) от 1 до 7, (b) от 8 до 14 и (с) от 15 до 21. Часть обследуемых получит список в последовательности (abc), другая – (bca), третья – (acb), четвертая – (cba), пятая – (bac) и шестая – (cab). Смещения ответов, связанные с различным уровнем внимания к началу и концу списка, будут погашаться по закону больших чисел.

Установлено, что у некоторых людей обнаруживается эффект монотонного «за» или «против» реагирования (*«галло эффект»*). Такие люди, ответив «да» на первый вопрос, отвечают «да» и на второй, и на третий, и так до 4–5 монотонных «да» или «нет» в случае ответов на однотипные вопросы. Это явление психологической ригидности. Чтобы уберечься от искажения данных такого рода, используют простой прием: *вопросы-«глушители»*. Перемежают однотипные вопросы другими, отличными от них по содержанию. Иногда для такой цели используют вовсе не нужные темы, единственное назначение которых – отвлечь внимание, устранить монотонность.

7. Ограничения количества выбора вариантов ответа могут быть жесткими и нежесткими. Это зависит от программной цели вопроса и его смысла. Если по смыслу вопроса возможны комбинации разных выборов, притом в любом количестве, нельзя без особых пояснений ограничивать выбор условием: «Укажите не более трех наиболее важных пунктов», например, при перечислении возможных занятий в свободное время. Однако в том же случае, если цель вопроса – выявить доминанту интереса, отношения, оценки и т.п., ограничить выбор вполне целесообразно: «Хотя у Вас, вероятно, не одно и не два любимых занятия в свободное время, просим в предложенном перечне указать не более трех наиболее привлекательных». Обратите внимание на то, что исследователь объясняет ограничения выбора. В противном случае респондент окажется в затруднении или же вовсе не станет отвечать на вопрос («Они полагают, что у меня только три любимых занятия?»).

8. Важную роль выполняет вариант, предполагающий *возможность уклониться от ответа* на закрытый вопрос: «трудно сказать», «затрудняюсь ответить», «не помню», «не знаю».

Формула уклонения от ответа подчеркивает, что респонденту предоставляют достаточную свободу. Это побуждает его более добросовестно относиться к опросу в целом. Замечено, что отсутствие такой формулы там, где она явно предполагается содержанием вопроса, повышает процент вообще уклоняющихся от участия в опросе.

В целом, сопоставляя возможности открытых и закрытых вопросов, можно сказать, что при пилотажном изучении проблемы, когда исследователю трудно предусмотреть все варианты ответов на вопросы, лучше включать в анкету открытые вопросы. На последующих этапах исследования большинство вопросов следует делать закрытыми и полужакрытыми вопросами.

*Прямые и косвенные вопросы.* В прямом варианте вопроса предусмотрен ответ, который следует понимать в том же смысле, как его понимает опрашиваемый. Ответ на косвенный вопрос предполагает расшифровку в ином, скрытом от респондента смысле.

*Прямой вопрос:* Если Вас не удовлетворяют условия труда, не могли бы Вы указать, что именно? (отметьте соответствующие пункты).

- Организация работы
- Взаимоотношения в рабочем коллективе
- Отношения с руководством
- Неинтересная работа
- Неудобная сменность

*Косвенный вопрос* ставится в случае, если затронуты проблемы, по которым опрашиваемые не склонны высказываться откровенно. Способы перевода из прямой в косвенную форму зависят от содержания темы. Наиболее распространенный способ замены прямых вопросов косвенными – перевод из личной формы в безличную.

*Личные и безличные вопросы* в равной мере относятся к оценкам и суждениям самого опрашиваемого, но во втором случае оценки имеют косвенный характер. Так, вместо личного прямого вопроса: «Как Вы считаете» – задают косвенный, безличный: «Некоторые полагают, что... Какие суждения, по Вашему мнению, наиболее справедливы?». Ожидается, что опрашиваемый выберет те суждения, которых он сам придерживается.

Безличная и полубезличная форма вопроса употребляется для выявления мнений, расходящихся с общепринятыми. В вариантах ответов подчеркивается, что все они возможны и опрашиваемый не будет выглядеть «белой вороной», если согласится с каким-то суждением.

### **Основные принципы построения анкеты**

1. Программная логика вопросов не должна быть смешиваема с логикой построения анкеты. Опросный лист строится под углом зрения психологии восприятия опрашиваемого. Учет особенностей восприятия респондентом текста анкеты – ведущий принцип, из которого следуют и все другие требования к ее построению.

2. Учет специфики культуры и практического опыта опрашиваемой аудитории как с точки зрения стилистики формулировки вопросов, так и с точки зрения общей структуры опросного листа. Например, при опросах рабочих вряд ли разумно пространно объяснять научные цели проводимой работы. Лучше подчеркнуть ее практическую значимость. Опрашивая же экспертов, следует указать и практические, и научные цели исследования.

3. Одни и те же вопросы, расположенные в разной последовательности, дадут разную информацию. Например, если сначала поставить вопрос об уровне удовлетворенности какой-то деятельностью и ее условиями (труда, быта и т.п.), а затем – вопросы на оценку частных особенностей деятельности (удовлетворенность содержанием работы, заработком, бытовым обслуживанием и прочее), то общие оценки будут влиять на частные, снижая (или, напротив, повышая) их независимо от специфики того или иного аспекта общей ситуации. Наблюдается, с одной стороны, стремление респондента психологически оправдать общую оценку и, с другой стороны, усиленное действие эффекта «эха» (галло-эффект), т.е. многократного повторения одной и той же оценки, отнесенной к общей группе проблем.

В таком случае следует частные вопросы ставить первыми, обобщающий – в конце соответствующего «блока», предваряя фразой: «А теперь просим Вас оценить в целом, в какой мере Вы удовлетворены ... своей нынешней работой... условиями жизни» и т.д. Оценка частных условий труда, быта и прочее предваряет общую, заставляет респондента более ответственно подойти к итоговой оценке, помогает разобраться в собственных настроениях [109].

4. Смысловые «блоки» опросного листа должны быть примерно одного объема. Доминирование какого-то «блока» неизбежно сказывается на качестве ответов по другим смысловым «блокам».

5. Распределять вопросы по степени трудности следует следующим образом. Первые вопросы должны быть более простыми, далее следуют более сложные (желательно событийные, не оценочные), затем – еще сложнее (мотивационные), потом – спад (сно-

ва событийные, фактологические) и в конце - наиболее сложные вопросы (один-два), после чего – завершающая «паспортичка» (данные, которые испытуемый сообщает о себе).

### **Последовательность смысловых разделов анкеты.**

А. *Введение*, в котором указано: кто (организация или научное учреждение) и для чего проводит опрос, как будут использованы данные; если требуется по содержанию вопросов – гарантия анонимности информации, инструкция по заполнению анкеты и способе ее возврата.

Б. *Вступительные вопросы* выполняют две функции: заинтересовать респондента и максимально облегчить ему включение в работу. Поэтому в начале текста не должны присутствовать трудные или беспокоящие вопросы. Наиболее удобны для этой цели вопросы сугубо событийного содержания. Трудные вопросы, поставленные в начале, могут отпугнуть, и это приведет к отказу участвовать в опросе. Открытые вопросы, требующие пространных комментариев, располагают ближе к середине анкеты; как контроль, они разрешаются и в конце, но не более одного-двух. Наиболее острые или интимные вопросы располагают в последней трети листа.

В. *Заключительные вопросы* по содержанию темы должны быть относительно нетрудными, так как надо учесть, что, работая с анкетой, люди постепенно утомляются. Здесь хороши шкалы оценок и другая информация в закрытых вариантах.

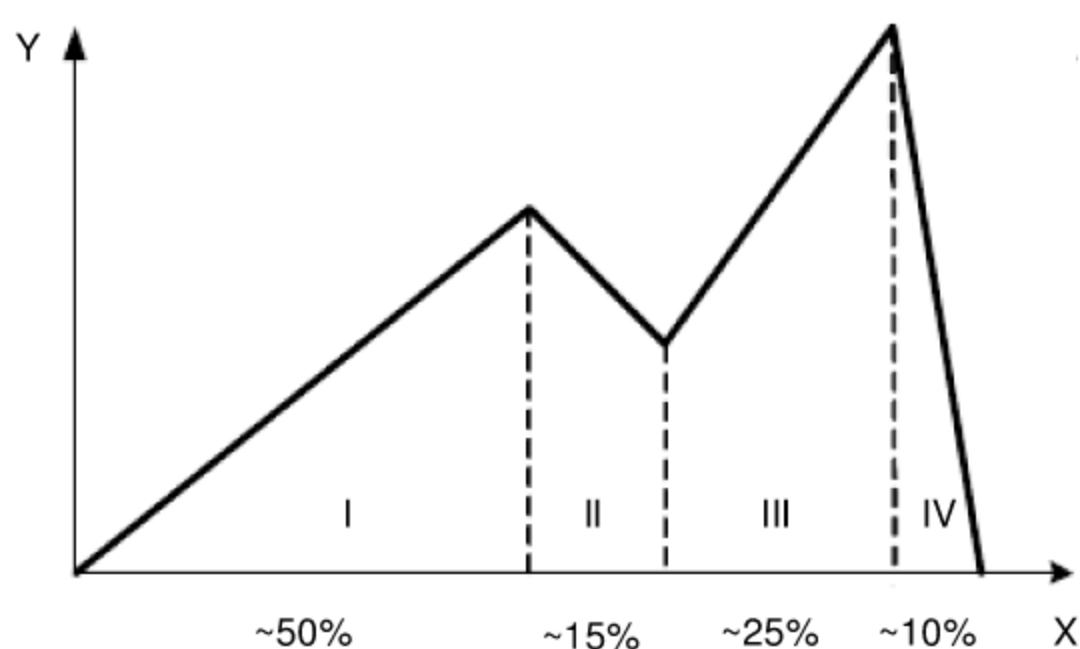
Г. «Паспортичка» (демографические сведения, которые испытуемый сообщает о себе: пол, возраст и т.д.) является последним смысловым блоком анкеты. Она лаконична, не требует особого напряжения и свидетельствует о завершении опроса.

Д. Обычно в заключение выражается благодарность за сотрудничество в проведении опроса. Часто это повторная благодарность, так как во введении пишут: «Заранее благодарим Вас за сотрудничество».

Общая продолжительность анкетного опроса составляет от десяти-пятнадцати минут до полутора-двух часов. *Динамика разви-*



тия анкетного опроса, представлена на рис. 4.1: в первой половине опроса – плавный подъем, примерно 15% времени уделено «отдыху» (спад), затем около четверти времени самой напряженной работы (к этому моменту респондент включился в опрос и подготовлен к серьезному обдумыванию своих ответов) и резкое снижение трудности вопросов в завершающей фазе.



**Рис. 4.1. Последовательность стадий постановки вопросов по уровню их сложности (Y) в процентах к общей длительности опроса (X) : I – плавный подъём; II – плавный спад; III – резкий подъём; IV – резкий спад (по В.А.Ядову)**

Верстка анкеты должна отвечать требованиям простоты и удобства работы и для опрашиваемого, и для кодировщика. Номер вариантов ответа должен включать номер вопроса и через точку – порядковый номер варианта ответа. Такая кодировка вопросов и вариантов ответа присваивает уникальный номер каждому пункту ответа, легка для восприятия респондентов и защищает от ошибок при оцифровке.

### **Пилотажная апробация анкеты**

При проведении пилотажной апробации анкеты исследователь сначала тщательно анализирует сконструированный опросный лист по всем тем критериям, представленным выше, затем размножает его в 25–30 экземплярах, которые и испытываются на опытной группе респондентов. Эта группа представляет микромодель пла-

нируемой выборки с выделением крайних значений ее параметров: полярных уровней образования, мужчин и женщин, других особенностей, существенных для представительности выборки. Правила подбора выборки для осуществления пилотажной апробации анкеты идентичны правилам подбора выборки стандартизации в ходе разработки строгоформализованной психодиагностической методики.

В углубленном «пилотаже», помимо этого, разрабатывают специальный путеводитель интервью для анкетеров, проводящих пробу. В нем предусматривают вопросы о понимании вопросов анкеты («интервью об интервью»), модификации конструкций закрытых вопросов, экспериментальные варианты формулировок одного и того же вопроса (в одной части опытного тиража даны одни, в другой – иные варианты), апробируют разные способы общей конструкции анкеты, варианты текста вводной части и т.п. Кроме того, проводящие пробу ведут наблюдение за поведением респондентов и их реакцией на вопросы, записывают и анализируют комментарии, которыми респонденты нередко сопровождают свои ответы, учитывают обстановку опроса и возможные факторы, мешающие получению адекватных ответов [109].

В ходе пилотажного обследования необходимо решить следующие задачи, сформулированные в виде вопросов:

1. Выдержаны ли требования к языку опрашиваемого, не получилось ли так, что для части респондентов язык слишком труден, для другой, наоборот, примитивен?

2. Все ли вопросы и варианты ответов понятны?

3. Не слишком ли абстрактны вопросы, или не слишком ли они конкретны?

4. Понятны ли респонденту единицы измерения, имеющиеся в анкете?

5. Компетентны ли опрашиваемые для ответов на вопросы, не следует ли включить «фильтры» на компетентность?

6. Нет ли опасности утомить опрашиваемого, как этого избежать, как снизить монотонность?

7. Достаточно ли надежна память опрашиваемого для ответов на вопросы о прошедших событиях, не следует ли подстраховаться на этот случай?

8. Нет ли опасности получения «угодных» или стереотипных ответов?

9. Не слишком ли многочисленны варианты ответов на вопрос, смогут ли респонденты справиться с обилием вариантов, как сократить их число или как расчленить их по блокам?

10. Нет ли опасений вызвать недоверие или какие-нибудь отрицательные эмоции у опрашиваемых?

11. Не слишком ли задевается самолюбие или интимные стороны жизни опрашиваемого?

12. Какие ответы следует истолковывать буквально, а какие - понимать не в прямом смысле, и в каком именно, как осуществить контроль на интерпретацию ответа?

13. Все ли в порядке со стороны графического оформления опросного листа, нет ли переносов в смысловых кусках, насколько четко выделены шрифтами вопросы и инструкция для ответа, нет ли монотонности, однообразия в оформлении?

14. Какие пункты следовало бы особенно тщательно проверить в пробе, нельзя ли сформулировать для них альтернативы и проверить наряду с имеющимся вариантом?

В ходе пилотажного исследования можно выявить следующие типичные недостатки по их внешним признакам [109]:

- нелогичность и пропуски в ответах, отсутствие порядка в комментариях на открытые вопросы: исследователь не учел различия в уровне культуры, компетентности и образования опрашиваемых;
- ответы типа «все или ничего», т.е. при многочисленных вариантах ответы группируются только в одном месте: результат стереотипных формулировок или неравного членения содержания ответов на предложенный вопрос; один вариант включает в себя остальные. Следует переформулировать и вопросы и варианты ответов;

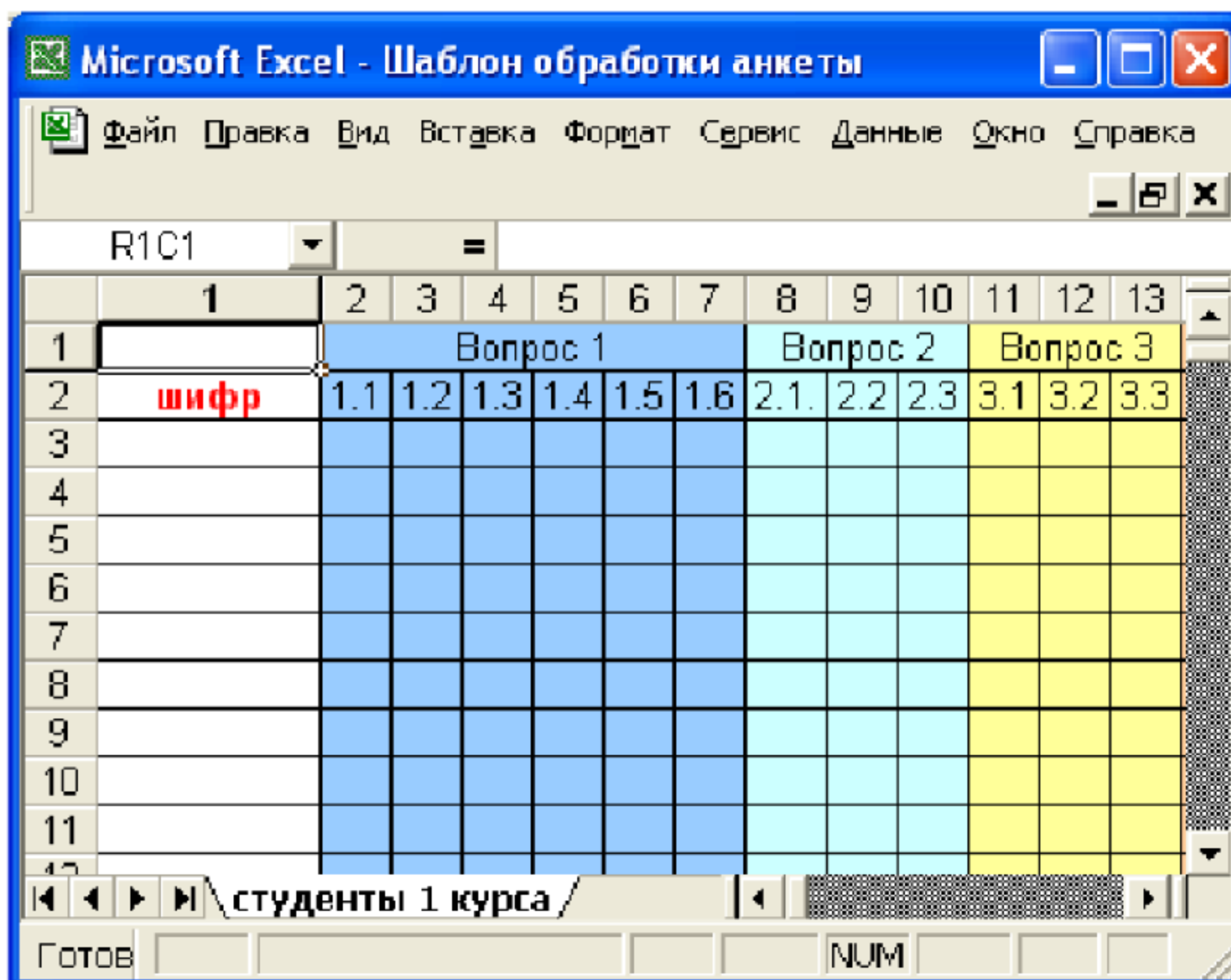
- большой процент отвечающих «не знаю», «не могу сказать», «не понял» (превышает 5–7% численности опрошенных) говорит о том, что: (а) вопрос или варианты ответов на него туманны, непонятны; (б) слишком сложны для аудитории, не отвечают ее опыту и знаниям; (в) суждения не расчленены на более простые составляющие; (г) единицы счета (если они есть) непонятны или необычны;
- множество дополнительных комментариев и замечаний к вопросам, где они не предусмотрены: в закрытых вопросах не полностью развернуто содержание возможных ответов. Дополнительные замечания опрошенных в пробе надо использовать для того, чтобы при массовом опросе более полно раскрыть содержание темы;
- существенные изменения в содержании ответов, если порядок вопросов и вариантов ответов на них изменяется. Следует в массовом опросе применять технику «блокирования» и перестановки блоков для разных подвыборок опрашиваемых;
- наконец, чрезмерная численность вовсе уклоняющихся от участия в опросе (больше 5%) свидетельствует: о нетактичности, необщительности интервьюера, неудачности общего плана опроса (трудные вопросы расположены близко к началу) или о том, что неблагоприятна сама ситуация опроса: респондента вынуждают прервать занятия своими делами, он торопится или взбудоражен предшествующими событиями, в помещении шумно либо присутствуют посторонние и т.д.

### **Оцифровка и первичный анализ результатов анкетного опроса**

В настоящее время существуют различные специальные сканеры, производящие перевод в цифровой формат данных с анкетного листа. При проведении оцифровки данных вручную удобно пользоваться электронными таблицами Excel.

Первоначально следует подготовить шаблон обработки анкет, указав в верхней строчке таблицы номера вопросов и варианты от-

ветов. Удобно залить определённым цветом столбцы, относящиеся к одному вопросу. Пример шаблона представлен на рис. 4.2



**Рис. 4.2** Пример шаблона обработки анкет в программе Excel.

В процессе обработки заполняются все ячейки. Объединение ячеек недопустимо. В ячейку ставится 1, если респондент отметил соответствующий вариант ответа. В противном случае в ячейку ставится 0. Если вопрос содержал открытую форму ответа, весь ответ вносится в ячейку.

После того, как таблица заполнена, для того, чтобы посчитать частоту тех или иных ответов в выборке, достаточно просуммировать значения данного столбца.

Дальнейший анализ данных проводится с использованием методов первичной математической обработки и статистического анализа, которые изложены в главах 6-8.

### Контрольные вопросы и задания

1. Чем обеспечивается достоверность данных, получаемых с помощью анкетного опроса?
2. Перечислите типы вопросов.
3. Каковы основные принципы построения анкеты?
4. Какова последовательность смысловых разделов анкеты?
5. Перечислите цели и содержание пилотажной апробации анкеты.

### 4.3. Беседа, интервью

Беседа и интервью относятся к методам сбора первичных данных на основе вербальной коммуникации. Интервью – это беседа, проводимая по определённому плану. В зависимости от целевого назначения интервью разделяют на диагностические и клинические.

**Диагностические интервью** – это метод получения информации общего содержания, направленный на «зондирование» различных аспектов поведения, свойств личности, характера. Оно может быть управляемым и неуправляемым («исповедальным»).

**Клиническое интервью** – это метод терапевтической беседы, помогающей человеку осознать свои внутренние затруднения, конфликты, скрытые мотивы поведения.

По форме интервью бывает:

- свободным (беседа без строгой детализации вопросов, но по общей программе (стойкая стратегия, свободная тактика);
- стандартизированным (с детальной разработкой всей процедуры, включая общий план беседы, последовательность вопросов, варианты возможных ответов: стойкая стратегия и тактика);
- частично стандартизированным (стойкая стратегия, тактика более свободная).

Диагностическим целям в большей мере соответствует стандартизированная форма интервью, поскольку она даёт возможность

- получить сравнимые данные по разным испытуемым;
- ограничить влияние посторонних воздействий;



- позволяет в полной мере и в нужной последовательности отрабатывать все вопросы.

Стандартизированное интервью практически ничем не отличается от опроса по анкете, за исключением того, что ответы записываются не самим респондентом, но интервьюером. К подобному способу прибегают для того, чтобы: (а) получить непосредственное впечатление от живой реакции опрашиваемых по предмету исследования (это помогает лучше интерпретировать их суждения) и (б) в случае, когда письменный опрос оказывается невозможным или затруднительным вследствие разнородности аудитории, необходимости пояснить многие вопросы с учетом различий в культуре и образовании респондентов и т.п.

Преимущества интервью перед анкетным опросом раскрываются в полной мере при использовании *частично стандартизированного или свободного интервью*. В таких интервью предусмотрен лишь список основных вопросов, частично их порядок (он может меняться по обстоятельствам), а получаемая информация служит для формулировки гипотез, выявления социальных проблем, подлежащих далее более систематическому анализу.

Прямой контакт с опрашиваемым и психологические отношения, которые устанавливаются между интервьюером и респондентом, создают немало преимуществ для получения информации, малодоступной путем анкетного опроса. Но существуют и недостатки, связанные с негативным восприятием респондентом личности интервьюера. В первую очередь, это касается интервьюеров, проводящие массовые опросы.

Задача интервьюера состоит в том, чтобы максимально непредвзято и объективно зарегистрировать ответы респондента на планируемые вопросы, ставить эти вопросы тактично, ровно, ненавязчиво, проявляя находчивость, быстроту реакции и умение «проигрывать» многообразные гипотезы относительно поведения респондента.

Темп речи интервьюера влияет на поведение респондента. Если интервьюируемый привык говорить быстро, медленный темп речи

интервьюера будет его раздражать. На ход беседы влияют соотношение в возрасте и пол участников разговора. Интервьюер примерно того же возраста, что и опрашиваемый, но противоположного пола обычно добивается лучшего результата.

Обстановка, в которой проходит беседа, должна располагать к спокойному и откровенному разговору. Нельзя проводить интервью в людных помещениях, при посторонних или в официальных кабинетах.

Интервьюер должен хорошо представлять себе общие цели исследования, его замысел, быть общителен по характеру (поэтому не каждый человек способен стать хорошим интервьюером), социально активен, обладать достаточно высокой культурой и образованием.

Стандартизированное интервью следует применять только тогда, когда с респондентом установлен контакт и он охотно отвечает на вопросы. В противном случае результат может быть неудовлетворительным, поскольку стандартизированное интервью многими людьми воспринимается как ситуация экзаменационного опроса, что ограничивает проявления непосредственности и искренности отвечающего. Интервью не должно быть продолжительным и скучным. Регистрация ответов не должна сдерживать отвечающего и может производиться по ходу разговора с разрешения интервьюируемого. Лучше всего, если беседу регистрирует ассистент интервьюера, который садится так, чтобы опрашиваемый мог видеть его боковым зрением, тогда как интервьюер располагается прямо напротив респондента. Этим достигается двойная цель: внимание приковано к интервьюеру и отвлечено от ведущего протокол. В то же время сам факт ведения протокола не скрывается, и это перестает волновать респондента. С разрешения респондента можно использовать диктофон.

Проводя беседу или интервью, очень важно занять правильную позицию по отношению к клиенту. Больше всего здесь подходит принцип недирективной психотерапии:

- создать атмосферу полного понимания, позволяющую как можно раньше установить контакт;
- принимать клиента таким, какой он есть;
- тактично и бережно относиться к позиции клиента, всё понимая, ничего не осуждая, но и не оправдывая.

Примерный план ведения беседы может быть следующим [109]:

### **1. Установление первого контакта.**

Вначале интервьюер называет себя и представляемую им организацию, помня, что не надо подчеркивать свою личную заинтересованность в содержании интервью, например: «Я – представитель такой-то организации. Меня зовут... Наша организация изучает условия обучения студентов в нашем вузе. Для этого мы хотели бы собрать некоторые сведения у студентов. - Вы не возражаете, если я задам Вам несколько вопросов?».

### **2. Закрепление контакта и первые вопросы по плану интервью.**

На этом этапе продолжается *общая разведка*. Как и в анкетных опросах, первые сведения - чисто фактуальные (обычные обязанности, описание условий жизни, быта). Сомнения в компетентности опрашиваемого и другие настораживающие вопросы на первом этапе недопустимы.

**3. Переход к основным вопросам** интервью должен сопровождаться вводными словами, которые подчеркивают важность последующего разговора. «Теперь позвольте перейти к некоторым вопросам, которые касаются условий Вашей учёбы».

После каждого ответа на основные вопросы следует невербально прореагировать на него: внимательный взгляд, одобрительный кивок, поддакивание. Допустимо частичное несогласие с опрашиваемым: «Вы говорите, что... Однако другие студенты полагают иначе...». Такое несогласие может послужить катализатором большей откровенности.

**4. Важный элемент искусства интервью – быстрое восстановление контакта** с респондентом в случае его утраты. Опрашивае-

мый может почему-то отказаться отвечать на вопрос или начинает отвечать невпопад. В этом случае возможно переспросить то же самое другими словами, либо напомнить о каком-то событии, либо плавно перейти к следующей теме.

**5. Завершение беседы.** В ходе беседы интервьюер подытоживает логические части беседы. В заключение он может вернуться к некоторым вопросам, на которые получены неполные ответы, и просит кое-что уточнить, ссылаясь на то, что теперь это кажется ему более важным, чем представлялось в ходе разговора.

Когда содержание интервью исчерпано, опрашиваемого просят дать некоторые сведения о себе, подчеркивая, что это надо для общей обработки данных.

Интервьюер благодарит за беседу, еще раз подчеркивает, что она была очень важна для исследования и что сведения, которые он записал, не будут использованы ни в каких иных целях, кроме изучения вопроса в целом.

Иногда опрашиваемый интересуется, что все-таки получится из этого исследования, принесет ли оно практически пользу ему и его сокурсникам. Интервьюер ни в коем случае не должен раздавать обещания и фантазировать насчет конечных итогов работы. Он указывает лишь общую цель исследования и ее практическую значимость.

Если респондент в ответ на поставленный вопрос интересуется, каково мнение самого интервьюера по этому поводу, последний говорит, что выскажет свои соображения в конце беседы: «Сейчас я не хотел бы сообщать, что думаю по этому поводу, это ведь может изменить ход нашей беседы. Давайте продолжим, а потом вернемся к этим вопросам».

Интервьюер может обещать информировать опрашиваемого о результатах обследования в целом, если тот обратится в организацию, адрес и телефон которой он предлагает записать.

**Оформление протокола интервью** производится на основе записи беседы тотчас или вскоре после ее окончания.

Следует максимально использовать выражения, слова, интонации речи опрашиваемого, излагать текст интервью не от третьего лица, а от первого.

Для оценки уровня контакта с респондентом фиксируют два показателя: длительность беседы, по мнению опрашиваемого (интуитивно) в сопоставлении с реальной длительностью и собственную оценку контакта, даваемую интервьюером. Например, опрашиваемый сказал, что беседа продолжалась 40 минут, тогда как реально она длилась около 60 минут: респондент чувствовал себя непринужденно. Это должно подтвердиться в субъективном впечатлении интервьюера, оценивающего контакт по пятибалльной шкале.

Не только ответы респондента, но и все вопросы, замечания интервьюера полностью заносятся в протокол. Дополнительные вопросы и пояснения часто не фиксируются в бланке. Поэтому регистрация всех реплик интервьюера столь же важна, как и запись ответов опрашиваемого.

Удобнее всего оформлять протокол так, чтобы, оставляя справа небольшие поля для заметок о поведении, мимике и т.п., о невысказанной информации, основной текст беседы протоколировать «лесенкой»: вопросы записываются с самого начала строки, а ответы – с отступом 2 см. Такую запись легче анализировать.

Заметки на полях протокола служат иногда дополнительным и важным источником информации. Здесь отмечают краткими ремарками внешние экспрессивные особенности поведения собеседника. Например: «сильно взволнован», «оживлен», «молчание», «нервничает», «воодушевился», «хочет угадать правильный ответ», «пришлось прервать беседу - отвлекли» и т.д. Опытный интервьюер пользуется условными значками для подобных заметок, которые он расшифровывает при окончательном оформлении протокола.

В заключение протокола желательно фиксировать общее впечатление о собеседнике, для чего выделяется раздел *«Особые отметки. Общее впечатление»*. Интервьюер и его ассистент подписывают протокол, отмечают дату и место проведения интервью.

Протокол формализованного интервью почти не отличается от заполненного анкетного бланка, ибо все ответы кодируются по заготовленному стандарту. Краткие заметки интервьюера делаются в специально отведенных местах.

*Качество информации*, получаемой путем интервью, зависит от всех перечисленных в этом разделе факторов, а также от уровня ответственности и добросовестности интервьюера. Непременным правилом является выборочный контроль работы каждого интервьюера. С этой целью организаторы исследования либо проводят краткое повторное интервью с одним из 10 (или около того) ранее опрошенных, либо удостоверяются, что интервью действительно имело место в такое-то время. Использование контроля способствует уверенности авторов и организаторов исследования в добротности данных, понимания специфики поведения опрашиваемых, что существенно помогает правильно интерпретировать протоколы интервью при общей обработке и анализе данных.

### **Контрольные вопросы и задания**

1. *Какими бывают интервью по форме?*
2. *Каким должно быть поведение интервьюера для обеспечения качества и достоверности собранных данных?*
3. *Предложите способ эффективной, с Вашей точки зрения, фиксации ответов респондента, а также дополнительной вербальной и невербальной информации.*

### **4.4. Метод экспертных оценок**

Метод экспертной оценки – малоформализованный метод, опирающийся на мнение экспертов, хорошо знающих оцениваемое явление и способных дать ему достоверную оценку. Предполагает изучение и обобщение мнений всех участвующих экспертов. Широко применяется в психологии личности. В качестве экспертов могут выступать компетентные лица, хорошо знающие испытуе-



мых: воспитатели, классные руководители, руководители коллективов и пр. Оценки экспертные целесообразно проводить не в виде описания качественных проявлений свойств – это лучше сделать в последующей беседе с экспертами, а в виде количественной оценки выраженности этих свойств или элементов поведения. Эксперты должны фиксировать выраженность более или менее мелких, частных элементов поведения, однозначно понятных. Обобщение этих элементов – дело самого исследователя [76].

Экспертная оценка – диагностический метод измерения, с помощью которого качественные особенности психических явлений получают свое числовое выражение в форме количественных оценок. Суть экспертного оценивания состоит в формализованном выражении объективной информации о поведении и качествах личности в ряде субъективных мнений и оценок. Этот метод может быть широко использован для оценки качеств личности при составлении психологических характеристик.

Метод экспертных оценок – формализованная процедура сбора, анализа и интерпретации независимых суждений достаточного количества экспертов о степени выраженности каждого из подлежащих оценке психологических качеств студента со стороны хорошо знающих его преподавателей, товарищей по учебе, которые имели возможность общаться с ним в различных ситуациях. Метод экспертных оценок имеет самостоятельную диагностическую ценность и может включаться в систему методов изучения личности с учетом его реальных достоинств.

Основным преимуществом метода экспертной оценки является возможность создания интегративного «психологического портрета» испытуемого, описания его личностных особенностей как системы взаимосвязанных, взаимозависимых и компенсирующих друг друга качеств. Это делает реальным более качественный прогноз обучаемости, позволяет целенаправленно выбирать средства педагогического воздействия с учетом возможной реакции на них данной личности, открывает пути к определению перспектив ее психологического и профессионального становления и развития.

При этом объективность собираемой экспертной информации обеспечивается привлечением широкого круга экспертов (преподавателей, сокурсников), которые наблюдают проявление индивидуально-психологических особенностей студента на протяжении достаточного времени при непосредственном общении в различных типах ситуаций, возникающих в процессе его обучения в университете.

Процедура проведения экспертного опроса включает следующие этапы:

- определение цели получения экспертных оценок;
- формирование рабочей группы и подбор экспертов;
- проведение опроса экспертов;
- обработка и анализ результатов экспертного оценивания.

Метод экспертного оценивания психологических качеств дает возможность оперативно собирать диагностическую информацию, быстро фиксировать динамику развития личности. Его применение не требует длительной подготовки и поиска или разработки сложного исследовательского инструментария. Этот метод является достаточно гибким, он легко модифицируется при возникновении новых задач, а также пригоден для многократного использования.

Практика применения метода экспертной оценки показывает, что его диагностическая ценность может быть существенно повышена за счет правильного подбора различных категорий экспертов, стандартизации и усовершенствования процедуры экспертного оценивания, включения дополнительных способов статистической обработки результатов оценивания, применением новых критериев анализа и интерпретации. Важное значение имеет обоснование выбора критериев (признаков, параметров), подлежащих экспертной оценке, отвечающих конкретным целям изучения индивидуально-психологических особенностей студентов. Это позволяет, в частности, создать более точную психологическую модель выпускника университета.

Следует также отметить, что целый ряд профессионально важных качеств личности бывает чрезвычайно трудно, а иногда и невозможно, выявить и оценить с помощью строгоформализованных методик, а также метода анкетного опроса и анализа документации. В этих случаях только экспертный метод дает возможность получить необходимые данные.

Особенно продуктивно применение метода экспертных оценок в сочетании со строгоформализованными методиками. Данные экспертных оценок при этом должны использоваться как подтверждающие или опровергающие (ставящие под сомнение) достоверность сведений, полученных при тестировании и анкетировании, они должны быть проанализированы также в качестве дополняющих, уточняющих или объясняющих результаты изучения личности другими методами. Любое рассогласование результатов психодиагностики с экспертными оценками должно стать предметом специального анализа и интерпретации.

Необходимо отметить и положительный эффект участия преподавателей в проведении экспертного оценивания психологических свойств и качеств студентов. Он проявляется в повышении уровня их психологической компетентности, развитии педагогических способностей, таких как наблюдательность, рефлексия, способность объективного видения и понимания индивидуальности каждого студента, умение сопоставить свое мнение об обучаемом с мнением других преподавателей и самих студентов.

Следовательно, метод экспертных оценок можно рассматривать как важный элемент в целостной системе методов психодиагностического обследования, проводимого в целях изучения личности студента. Диагностическая сущность метода экспертной оценки, состоит в получении объективной информации об индивидуально-психологических особенностях, в результате специально организованного опроса мнений экспертов.

Однако, экспертные оценки не всегда являются достоверной мерой психологических качеств, так как восприятие эксперта под-

вержено субъективным искажениям в связи с особенностями его личности.

Поскольку различные эксперты будут давать различные оценки, возникает проблема их надежности. Одним из примеров субъективных ошибок является влияние на оценки хорошего или плохого отношения эксперта к оцениваемому студенту. Другим примером ошибок, вносимых в оценку качеств личности, являются инструментальные искажения, причиной которых может быть различие в способах измерения. Причем они наиболее значительны именно при методе экспертных оценок.

Повышения надежности экспертных оценок можно добиться за счет снятия субъективных и инструментальных искажений. Для этого обычная процедура сбора экспертных оценок должна быть дополнена специальным приемом повторного оценивания с предварительным сообщением эксперту данных, полученных от других экспертов. Изменение в этом случае первоначальной оценки является важным диагностическим показателем ненадежности эксперта. Кроме того, для повышения надежности метода экспертных оценок необходимо выполнять следующие требования:

- оцениваемые качества должны определяться в терминах наблюдаемого поведения;
- эксперт должен иметь возможность наблюдать за поведением оцениваемого студента достаточно длительный промежуток времени;
- необходимо не менее пяти экспертов на одного оцениваемого студента;
- ранжирование должно производиться экспертами только по одной черте за один раз, вместо оценивания одного испытуемого сразу по всему перечню критериев;
- данные экспертного оценивания необходимо соотносить с результатами объективных тестовых обследований.

### **Контрольные вопросы и задания**

1. *Охарактеризуйте метод экспертной оценки в структуре методов психологического исследования.*

2. *Перечислите этапы процедуры проведения анкетного опроса.*
3. *Какие требования необходимо выполнять для повышения надёжности метода экспертных оценок?*

#### **4.4. Контент-анализ**

Контент-анализ (content analysis; от content – содержание, analysis – анализ, изучение, исследование) – формализованный метод изучения текстовой и графической информации, заключающийся в переводе изучаемой информации в количественные показатели и ее статистической обработке [76]. Контент-анализ начал применяться с 20-х г.г. XX века для обработки средств массовой коммуникации и представляет собой формализованный метод анализа документов, когда на основе гипотезы исследователя в документальных материалах выделяются специальные единицы информации, а затем подсчитывается частота их употребления. В настоящее время метод применяется для анализа информации в области социальной психологии и социологии.

Количественно-качественный анализ продуктов деятельности является одним из способов получения социологической и психологической информации о человеке, коллективе, обществе. Под продуктами деятельности здесь понимаются личные документы (письма, автобиографии, дневники, фотографии), результаты анкетных опросов, материалы групповой, коллективной коммуникации (записи разговоров, дискуссий, совещаний, различные уставы, приказы, объявления), а также сочинения, литературные произведения, произведения искусства. Объектом контент-анализа традиционно является содержание материалов средств массовой информации (печатных изданий, радио- и телепередач), публичных выступлений, теле и кинофильмов, рекламных роликов и т.д.

В нашей стране количественные методы при изучении документов начали применяться с 20-х годов XX века, эти методы использовали в своих исследованиях Н. А. Рыбников, П. П. Блонский, И.Н.Шпильрейн, В.А.Кузьмичев и др.

Качественно-количественный анализ содержания в 20-е годы XX века использовал в своих работах известный русский советский исследователь биографических материалов Николай Александрович Рыбников [86]. В ходе исследований им был разработан прием побуждения испытуемого говорить о себе по определенному плану, гарантирующий «однообразие собираемого материала, что имеет огромные преимущества, давая возможность сравнивать, объединять, обобщать собираемые факты и т. д.» [86, с.40]. В 1926-1928 гг. Н. А. Рыбниковым были проанализированы более 500 автобиографических сочинений детей рабочих на тему: «Как я живу теперь». В ходе анализа сочинений было выявлено распределение положительных и отрицательных оценок школьниками своей жизни в зависимости от возраста и пола. В частности Н. А. Рыбниковым было установлено, что в младшем возрасте девочки дают более высокий процент положительных оценок по сравнению с мальчиками. У школьников старшего возраста получен противоположный результат, что связано с возрастанием нагрузки по работе в семье у девочек старшего возраста. Далее Рыбников анализирует «мотивы того или иного события, мотивы общего жизненного процесса, встречающиеся в детских автобиографиях». Эти мотивы он разбивает на следующие три группы: материальные, психологические и неопределенные. Чаще всего встречаются мотивы материального характера (53 %), мотивы психологического характера дает около одной трети ребят (31 %). Н.А.Рыбниковым установлено, что у детей рабочих «хорошее житье чаще всего мотивируется причинами психологического характера (67 % против 25 %); наоборот, плохое житьё-бытьё вдвое чаще обосновывается мотивами материального характера (31 % против 16 %)» [86, с.42].

Качественно-количественный метод исследования позволил отечественным психологам первой половины XX века получить результаты, ставящие под сомнение известные западные теории, в том числе психоаналитическую теорию З.Фрейда. Так, результаты исследования П. П. Блонского, в котором он проанализировал 190



собранных им «первых воспоминаний» учителей и студентов, а также 83 письменные работы школьников 11-13 лет на тему «Мое самое раннее воспоминание детства», привели его к заключению, что содержанием 68 % воспоминаний взрослых и 74 % воспоминаний школьников является несчастье. «Несчастье и страх – таковы основные мнемонические факторы», что противоречит фрейдовской теории забывания как вытеснения неприятного [цит. по 61, с.73]. К аналогичным выводам относительно теории Фрейда приводят и результаты исследования В. Н. Касаткина, в котором на протяжении 30-50-х годов XX века проводилось медико-психодиагностическое изучение содержания 4000 записанных сновидений здоровых и больных людей. При этом учитывались основные особенности качеств и условий жизни людей, сновидения которых изучались: возраст, пол, образование, специальность, состояние здоровья, семейное положение, родной язык и владение другими языками, местожителство, биографические сведения, дата, содержание дня, предшествующего сновидению, и состояние испытуемого при пробуждении. Автором было установлено, что в «сновидениях взрослых людей встречались элементы, связанные с работой, трудовой деятельностью (специальностью), в 62,5 % всех сновидений; элементы быта, как то: жилище, одежда, пища и т. п. - в 41,4 % всех сновидений; элементы, связанные со здоровьем, – в 44,3 %; эпизоды из семейной жизни – в 38,6 %, сексуальные - в 8,0 % всех сновидений» [там же, с.74]. Полученные данные позволили В. Н. Касаткину оспаривать теорию сновидений З.Фрейда.

Впечатляющие результаты отечественных психологов, использовавших качественно-количественный анализ, как основной метод исследования, заложили основу формирования методологии эмпирических социальных исследований.

Детальное описание процедуры контент-анализа была разработана американскими социологами Х. Лассуэллом и Б. Берелсоном. К ним относятся:

1. **Выявление смысловых единиц анализа**, которыми могут служить: а) понятия, выраженные в отдельных терминах; б) темы,

выраженные в целых смысловых абзацах, частях текстов, статьях, радиопередачах и т. п.; в) имена, фамилии людей; г) события, факты и т. п.; д) смысл апелляций к потенциальному адресату. Единицы контент-анализа выделяются в зависимости от содержания, целей, задач и гипотез конкретного исследования.

**2. Выделение единиц счета**, которые могут совпадать либо не совпадать с единицами анализа. В первом случае процедура сводится к подсчету частоты упоминания выделенной смысловой единицы. Во втором случае исследователь на основе анализируемого материала и здравого смысла сам выдвигает единицы счета, которыми могут быть: а) физическая протяженность текстов; б) площадь текста, заполненная смысловыми единицами; в) число строк (абзацев, знаков, колонок текста); г) длительность трансляции по радио или ТВ; д) метраж пленки при аудио- и видеозаписях, е) количество рисунков с определенным содержанием, сюжетом и пр.

**3. Процедура подсчета** в общем виде сходна со стандартными приемами классификации по выделенным группировкам. Применяется составление специальных таблиц, применение компьютерных программ, специальных формул (например, «формула оценки удельного веса смысловых категорий в общем объеме текста»), статистические расчеты понятности и аттрактивности текста.

Американскими исследователями контент-анализ стал использоваться с 40-х годов и для определения психологических особенностей, психических состояний личности и групп. В.С.Аванесов в работе «Тесты в социологическом исследовании» [2] даёт анализ сфер применения контент-анализа в различного рода социально-психологических и социально-политических исследованиях, проведённых в США.

Так, психологическую структуру отдельной конкретной личности на основе анализа коллекции личностных документов исследовали Г. Оллпорт [110] и А. Болдуин [111]. Психическую напряженность, предсуицидные состояния и мотивацию посредством анализа содержания писем, записок, дневников пытались измерить Дж. Доллард и О. Маурер [112], Ч. Осгуд и Е.Уолкер [115].

В целях специфической психолого-политической диагностики изучали различия в социальных ценностях у представителей США и Германии К. Левин [113] и Х. Себалд [117]. Первый исследователь анализировал американскую и немецко-фашистскую литературу для юношества, а второй - песенники этих же стран, изданные в 1940 г. В обоих исследованиях были обнаружены явные различия в ценностных ориентациях, которые пропагандировались американскими и немецкими изданиями тех лет.

К подобного же рода анализу относится исследование использования эмоциональных стереотипов в газете «Чикаго Трибюн», оппозиционно настроенной к президенту Рузвельту и его политике, проведенное в 30-х годах С. Сарджентом [116]. Для выражения отношения к политике и практике рузвельтовского направления газета использовала негативные стереотипы типа «диктатура, инквизитор, регламентация, подачка», тогда как в газете «Нью-Йорк таймс» в аналогичных случаях употреблялись термины: «контроль, исследователь, регулирование, помощь». Множественное сопоставление соответствующих выражений и понятий в газетах выявило отношение стоящих за газетами групп к определенным политическим лидерам, партиям и явлениям.

Л. Лоуэнталь на основе количественного анализа биографий, публикуемых в популярных журналах, показал [114], как изменялись ценности и кумиры американского общества на протяжении первых четырех десятилетий нашего века от «идолов производства» (бизнесмены, менеджеры, банкиры и т. д.) к «идолам потребления» (певцы, кинозвезды, спортсмены и т. п.).

С 50-х годов получает распространение качественно-количественный анализ вербальной коммуникации в малых группах, начатый Р. Бейлсом (следует отметить, что обычно такие исследования принято относить к наблюдению, хотя речь, зафиксированная, например, на магнитной ленте, становится уже документом). Посредством анализа диалогов, деловых бесед, дискуссий в малых группах и первичных коллективах можно диагностировать стиль руководства, социально-психологический климат, конфликт-

ность и т.п. [59]. Аналогичные возможности открываются для психодиагностики процессов и состояний при изучении массового вербального поведения на улице, в транспорте, магазинах и т.д. [118].

Новым этапом в развитии контент-анализа начиная с 60-х годов XX в. связано с использованием специальных компьютерных программ для анализа текстов. В Массачусетском технологическом институте был разработан «универсальный анализатор» (The General Inquirer) – комплекс компьютерных программ анализа текстовых материалов, при помощи которого можно подсчитывать частоты категорий содержания текста, получать различные индексы на основе совместного появления этих категорий и т.д. [119]. С помощью этих программ были исследованы речи двадцати американских президентов при их вступлении на этот пост, редакционные статьи в газетах разных стран, личные письма, сочинения, вербальное поведение психически больных людей и прочие материалы. Компьютерный контент-анализ приобрёл популярность в связи со снижением трудоёмкости и повышением точности анализа текстов. Вместе с тем, существуют определённые трудности семантического плана в области формализации материалов, подвергающегося анализу. Как отмечает В.С.Аванесов [61], важно правильно определить, когда следует воспользоваться компьютерным, а когда ручным анализом, что зависит от задач исследования, от объема материалов, подлежащих анализу, от степени их формализуемости.

В настоящее время развития информационных технологий появились более широкие возможности для анализа текстовых материалов. При этом подсчитываются не только частоты категорий содержания текста, но и совместное появление этих категорий. В США подобным образом исследуются различные материалы, начиная с вербального поведения психически больных людей и заканчивая редакционными статьями в газетах разных стран и речами президентов при их вступлении на этот пост. Методика применения контент-анализа в социально-политической сфере напрямую связана с именем французского исследователя Ж.Кайзера.

В начале 60-х гг. Ж. Кайзер разработал систему изучения больших текстовых массивов, основанную на анализе статистических данных. Суть методики Ж. Кайзера заключалась в учете внешних форм организации текстового материала: его расположения, оглавления, оформления, проч. Кроме этого Ж. Кайзер предложил целый комплекс универсальных эмпирических процедур, позволяющих достаточно полно, скрупулезно анализировать как отдельные исторические источники, так и их совокупности. Дальнейшее развитие этот метод получил в трудах последовательницы Ж.Кайзера Э.Морен, которая ввела в научный оборот понятия «единицы информации» (особые элементы счета в контент-анализе, включающие в себя ответ на вопрос: «О чем говорится?»), «смысловые группы» (совокупности «единиц информации», формирующиеся на основе их взаимного смыслового соответствия), «эмоциональный тон повествования» (этот тон определялся по специальной порядковом шкале с делениями: «положительная», «отрицательная», «нейтральная»). Исследования Э.Морен положили начало более «гибкому» подходу к процедуре статистической обработки текстов. Этот метод позволяет объективно исследовать происходящие политические явления и процессы [61].

Исследование начинается с анализа текстов, в которых содержится информация об указанных явлениях и процессах. Прежде всего, выделяются соответствующие смысловые единицы: определенные сведения, количественные показатели, оценки, понятия, раскрывающие содержание того или иного политического текста и, следовательно, отражаемого им политического явления. Поскольку содержащиеся в тексте сведения, оценки, понятия выражаются определенными терминами и характерными словосочетаниями, они также учитываются при контент-анализе.

Таким образом, контент-анализ начинается с логической, лингвистической и другой формализации изучаемого текста (в данном случае текста политического содержания). Поскольку чаще всего исследуются большие массивы информации, порой весьма сложные, то выделяется множество смысловых единиц и разрабатыва-



ется соответствующий математический аппарат их количественного анализа. В настоящее время контент-анализ широко применяется в аналитических исследованиях, проводимых в интересах бизнеса, политической аналитики и практики, в том числе силовых структур. И хотя текущие публикации СМИ по-прежнему составляют основную часть материалов, которые прорабатываются с применением этой методики, все большее распространение получает ее применение для направленного анализа содержания конкретных политических документов, важнейших выступлений политических деятелей или отдельных тематических очерков.

При всем разнообразии подходов к классификации инвариантов контент-анализа в специальных публикациях обычно выделяют две основные формы применения методики: направленный и ненаправленный контент-анализ. Различают также количественный и качественный варианты. Количественный (частотный) контент-анализ в обязательном порядке включает стандартизированные процедуры подсчета выделенных категорий. Для формулирования итоговых заключений решающее значение имеют количественные величины, которые характеризуют ту или иную категорию. Например, если исследователь стремится получить представления о том, как велика значимость понятия «мирное урегулирование» для лидеров воюющих сторон, и согласен с допущением, что она приблизительно определяется частотой упоминания этого понятия в официальных речах, то тогда, после соответствующих арифметических подсчетов, он получит определенные количественные показатели. Показатели могут отличаться или, напротив, быть близки по абсолютному значению, которое будет учитываться при интерпретации результатов обработки. Задачу можно усложнить, поставив в качестве предварительного условия выделение всех содержательных в смысловом отношении единиц соответствующих текстов, а затем подсчитав относительную значимость интересующего нас выражения по сравнению с другими. Примечательно, что в обоих случаях основная часть подсчетов может быть выполнена с применением простых компьютерных программ.



Качественный контент-анализ нацелен на углубленное содержательное изучение текстового материала, в том числе, с точки зрения контекста, в котором представлены выделенные категории. Итоговые заключения формулируются здесь с учетом взаимосвязей содержательных элементов и их относительной значимости (рангом) в структуре текста. Так, для того чтобы сравнить отношение разных авторов к проблеме мирного урегулирования, исследователь должен стремиться не просто выделить соответствующее понятие, но и определить, является ли его проблематика главной в системе декларируемых высказываний.

Ко всем контент-аналитическим методикам предъявляется ряд общих методических требований: обоснованности, надежности, объективности и системности. Так, обоснованность устанавливается путем соответствия выделенных категорий или понятий действительным реалиям исследуемой политической ситуации. Надежность осуществляется благодаря получению сходных (с допустимым расхождением не более пяти процентов) результатов различными кодировщиками при кодировании одних и тех документов по единой инструкции. Объективность определяется соответствием выбора категорий и единиц анализа всестороннему исследованию политической ситуации. Системность достигается на основе достаточно широкого круга материалов, включаемых в информационную базу исследования.

Указанные требования не только должны неукоснительно соблюдаться при проведении прикладного исследования, но и аргументированно подтверждаться в специальных методических разделах соответствующих проектов. Без проработки этих вопросов результаты проведенного на основе методики контент-анализа исследования считаются некорректными и утрачивают свой познавательный смысл.

Существуют четкие методические требования к возможной единице анализа: она должна быть достаточно большой, чтобы выражать определенное смысловое значение; достаточно малой, чтобы не выражать много значений; свободно идентифицироваться. Кро-

ме того, число первичных единиц аналитического наблюдения (элементов) должно быть достаточно велико, чтобы служить основой для статистической выборки. Второй этап применения методики контент-анализа является декомпозиция текста: выделение из содержания множества смысловых элементов, последующее упорядочение которого зависит от типа или варианта методики. Выбрав смысловую единицу и ее признаки, исследователь должен определить также и единицу счета, которая станет использоваться для количественного анализа материала. Третий этап применения методики контент-анализа предусматривает саму процедуру подсчета или квантификацию множества элементов текстового массива. В общем виде она аналогична стандартным приемам квантификации по выделенным группировкам. Например, изучая тематику газеты, можно определить процент конкретных тем по тематическим единицам разного содержания (политическим, экономическим, военно-техническим вопросам). Существуют также специальные процедуры подсчета, приспособленные для контент-анализа, например формула коэффициента Яниса, предназначенная для исчисления соотношения между благоприятными и неблагоприятными (относительно принятой позиции) оценками, суждениями, аргументами. Кроме того, применяются и более тонкие способы количественного описания содержания текста. Они связаны с построением шкал для оценки выделенных характеристик текста.

Четвертый этап применения контент-анализа состоит в содержательной интерпретации результатов декомпозиции текстового массива и квантификации выделенных элементов. Обычно он построен на выявлении и оценке таких характеристик текстового материала, которые позволяют дать ответ на вопрос, что хочет подчеркнуть (скрыть) его автор и что является в его понимании приоритетным направлением политических действий. Содержательная интерпретация завершается формулированием аналитических выводов, которые соотносятся с целями и задачами исследования, декларированными на его начальном этапе. На заключительном пятом этапе применения методики обычно происходит верификация получен-

ных результатов, которая может проводиться либо путем повторения исследовательской процедуры другими исполнителями, либо на основе критического осмысления и сопоставления новых данных с уже имеющейся информацией. Полезным считается также и их соотнесение с информацией, полученной путем применения иных исследовательских подходов. Первый вариант процедуры верификации обычно применяется в рамках пилотных научных проектов, а другие чаще используются в рутинных исследованиях. В случае успешной верификации результатов аналитической работы они фиксируются в итоговом документе конкретного прикладного проекта и могут служить как самостоятельной, так и промежуточной основой для подготовки практических шагов в сфере политической практики.

Накопленный к настоящему моменту опыт применения качественно-количественного анализа документов открывает его значительные возможности, в том числе, и для психологической оценки, причем как на уровне личности, так и на уровне малых и больших групп. В качестве эмпирических объектов изучения могут быть использованы личные документы (письма, фотографии, дневники, автобиографии и т. п.), материалы групповой, коллективной и массовой коммуникации (записи разговоров, дискуссий, совещаний, всевозможные уставы, приказы, объявления, газеты, радиопередачи, реклама и т. п.), продукты деятельности людей, включая литературу и искусство, а также при анализе ответов на открытые вопросы анкет, материалов наблюдений, результатов психологического тестирования (в частности, в проективных методиках), для анализа результатов в методе фокус-групп. Контент-анализ может быть использован для исследования большей части документальных источников, однако лучше всего он работает при относительно большом количестве однопорядковых данных [76]. Помимо самостоятельного применения или равноправного применения в комплексе с другими методами контент-анализ может выступать и в качестве вспомогательной техники для обработки данных, полученных посредством проективных методик (например, ТАТ, тест

Роршаха), нестандартизованных интервью, открытых вопросов анкет и т. п. [43].

В методическом аспекте контент-анализ находит применение в трех качествах: как основной метод, направленный на получение наиболее важной информации об изучаемом явлении; как параллельный метод, применяемый в комплексе с другими; как вспомогательный метод или процедура обработки данных, полученных при других исследованиях. В последнем качестве контент-анализ наиболее часто используется в психологической диагностике [11, с. 134]. Источниками информации для контент-анализа могут служить различные виды речевой продукции, материалы и документы, прямо или косвенно характеризующие изучаемые особенности личности. В психологической диагностике контент-анализ используется для анализа речевых сообщений испытуемого в ходе беседы и при выполнении проективных методик, а также продуктов деятельности (сочинения, дневники, рисунки). Контент-анализ позволяет выявить скрытую тенденцию в ответах и других видах информации, характеризующих испытуемого, относительно точно регистрировать внешне неразличимые показатели в объемных массивах эмпирических данных.

Л.Ф.Бурлачук выделяет следующие основные направления применения контент-анализа при проведении психодиагностических процедур [11]:

- анализ результатов применения проективных методик исследования личности (качественно-количественный анализ «тематических» высказываний испытуемого в процессе психодиагностики);
- изучение диагностических интервью, содержания бесед, другой речевой и письменной продукции испытуемого;
- анализ содержания открытых вопросов при массовом анкетировании;
- изучение психологических особенностей личности по качественным характеристикам, развернутым экспертным оценкам;
- анализ объективной информации о личности;

- анализ высказываний при диагностике особенностей групповой коммуникации.

Проведение контент-анализа в любом указанных из направлений требует выполнения следующих требований к организации его этапов и процедур:

1. Определение категорий анализа – ключевых элементов, регистрируемых в соответствии с задачами исследования. Эти категории (к числу которых относятся количество речевых актов взаимодействий, побуждающих высказываний, частота отражения отдельных тем: «болезнь», беспокойство по поводу «карьеры», «семьи» и т. д.) должны быть исчерпывающими, охватывать все части фиксируемого содержания. В случае включения контент-анализа как элемента процедуры обработки данных теста критерий должен быть четко формализован и «стандартизирован» в целях однозначности его характеристики лицами, проводящими исследование.

2. Определение соответствующие категориям единицы анализа – лингвистические единицы речи или элементы содержания, экспрессивные элементы речевой продукции (слова, суждения, темы, интонации, описываемые или наблюдаемые ситуации).

3. Выделение единиц квантификации (частоты появления анализируемых единиц в определенном объеме информации). Эти исходные данные составляют инструкцию по обработке – алгоритм кодирования.

Наиболее существенными ограничениями при выборе источников информации для контент-анализа является частота встречаемости необходимых данных и возможность их квантификации. Контент-анализ применяется только тогда, когда имеется достаточное количество материала для анализа (представлено много отдельных однородных документов, автобиографий, писем, фотографий и т. д. или есть несколько и даже один документ, например дневник, но достаточного объема). При этом интересующие элементы содержания (единицы анализа) встречаются в исследуемых документах с достаточной частотой. В противном случае не обеспечивается ста-



тистическая достоверность выводов. Критерием здесь служит закон больших чисел.

Не все виды документов пригодны для контент-анализа по причине затруднений с формализацией их содержания. Иногда невозможно задать четкие однозначные правила для фиксирования нужных характеристик содержания.

Существуют специальные процедуры подсчета, приспособленные для контент-анализа, например формула коэффициента Яниса, предназначенная для исчисления соотношения между благоприятными и неблагоприятными (относительно принятой позиции) оценками, суждениями, аргументами. Представляет интерес также метод Ч. Осгуда, направленный на выявление случайных и неслучайных элементов содержания. Этот вид анализа включает подсчет числа элементов каждого типа и числа совместных их наблюдений. На основании этих исходных данных вычисляется математически ожидаемая вероятность совместного появления:

$$P_{ab} = P_a * P_b$$

Если обнаруженное число совместных появлений единиц существенно больше, чем ожидаемое, то можно предположить, что возникшие сочетания не случайны. Установленная таким образом гипотеза может быть проверена с помощью других статистических и эмпирических методов.

Контент-анализ часто используется в качестве вспомогательного метода или процедуры обработки данных, полученных при других методах исследования. С его помощью подвергаются анализу речевые сообщения испытуемого (например, полученные в ходе нестандартизированной беседы), открытых вопросов анкет, он применяется при обработке данных, полученных посредством проективных методик (ТАТ, тест Роршаха, методика «Завершение предложений»).

В методиках диагностики личностных особенностей проводится контент-анализ грамматических и стилистических конструкций речи испытуемого: количество «тематических» высказываний (бо-



лезнь, страх, неуверенность и т.д.), глаголов, и т.п. Такой анализ позволяет выявить скрытые тенденции в ответах испытуемого.

### **Контрольные вопросы и задания**

1. *Перечислите сферы применения контент-анализа.*
2. *Изложите историю развития качественно-количественного анализа информации в нашей стране и за рубежом.*
3. *Опишите суть качественного и количественного анализа информации*
4. *Перечислите основные направления применения контент-анализа при проведении психологической оценки.*

## **Глава 5.**

### **Психологическая оценка компетенций в образовательном процессе**

#### **5.1. Компетентностный подход к результатам высшего профессионального образования**

Федеральные образовательные стандарты высшего образования (ФГОС ВО) представляют собой совокупность требований, обязательных при реализации основных образовательных программ бакалавриата, специалитета и магистратуры по различным направлениям подготовки. По каждому направлению даётся полная характеристика профессиональной деятельности, включающая область, объекты, виды, задачи профессиональной деятельности, а также представлены требования к результатам освоения образовательных программ, в соответствии с которыми выпускник должен обладать определённым набором общекультурных и профессиональных компетенций. Знания, умения, навыки и освоенные компетенции являются связующим звеном между ступенями непрерывного образования, а также между результатами образования и требованиями работодателя.

Если знания студентов и выпускников высших учебных заведений подлежат оценке традиционным способом в форме зачётов и экзаменов, то каковы же методы оценки общекультурных и профессиональных компетенций будущих специалистов и кто может проводить эту оценку? Для ответа на этот вопрос рассмотрим содержание компетентностного подхода к результатам образования.

В настоящее время компетентностный подход является одним из доминирующих факторов модернизации российского образования и связан с повышением интереса современного общества к психологическим ресурсам людей. Понятие компетентности, по

словам М. А. Холодной, – это пример понятия, отвечающего требованиям экологического подхода в психологических исследованиях человека [101]. Как отмечает А. М. Новиков [57, с.333], компетентностный подход основан на концепции компетенций как основе формирования у обучающихся способностей решать практические задачи и воспитания личности в целом. Компетентность, по определению А. М. Новикова – это самостоятельно реализуемая способность к практической деятельности, к решению жизненных проблем, основанная на приобретенных обучающимся учебном и жизненном опыте, его ценностях и склонностях.

Компетентностный подход в образовании, когда целью образовательного процесса является приобретение и обогащение личностно- и мотивационно-обусловленного опыта социально-профессиональной жизнедеятельности человека начал складываться с 60-х годов XX века. Дж. Равен определяет компетентность как специфическую способность, необходимую для эффективного выполнения конкретного действия в конкретной предметной области и включающую узкоспециальные знания, особого рода предметные навыки, способы мышления, а также понимание ответственности за свои действия. По Дж. Равену существуют также «высшие компетентности», которые вне зависимости от того, в какой сфере они проявляются, предполагают наличие у человека высокого уровня инициативы, способности организовать других людей для достижения поставленных целей, готовности оценивать и анализировать социальные последствия своих действий и т.д. В своей работе «Компетентность в современном обществе» Дж. Равен рассматривает 39 видов компетенций [79, с.281-296], которые он определяет как «мотивированные способности» [там же, с.258], относящиеся к различным сферам психики: когнитивной, эмоциональной, мотивационной, личностной (табл. 5.1).

**Таблица 5.1 Структура компетенций по Дж. Равену**

Сфера компетентности	Виды компетенций по Дж. Равену
Когнитивная	готовность и способность обучаться самостоятельно; готовность решать сложные вопросы; готовность работать над чем-либо спорным и вызывающим беспокойство; знание того, как использовать инновации; самостоятельность мышления, оригинальность; критическое мышление; исследование окружающей среды для выявления ее возможностей и ресурсов (как материальных, так и человеческих); внимание к проблемам, связанным с достижением поставленных целей; готовность использовать новые идеи и инновации для достижения цели; поиск и использование обратной связи; склонность к размышлениям о будущем: привычка к абстрагированию
Ценностная	тенденция к более ясному пониманию ценностей и установок по отношению к конкретной цели; отношение к правилам как указателям желательных способов поведения
Эмоциональная	вовлечение эмоций в процесс деятельности
Личностная	уверенность в себе; самоконтроль; адаптивность: отсутствие чувства беспомощности, готовность полагаться на субъективные оценки и идти на умеренный риск; отсутствие фатализма; уверенность в благожелательном отношении общества к инновациям; доверие; способность принимать решения; персональная ответственность; настойчивость; тенденция контролировать свою деятельность; использование ресурсов
Коммуникативная	способность к совместной работе ради достижения цели; способность побуждать других людей работать сообща ради достижения поставленной цели; способность слушать других людей и принимать во внимание то, что они говорят; установка на взаимный выигрыш и широта перспектив; стремление к субъективной оценке личностного потенциала сотрудников; готовность разрешать другим людям принимать самостоятельные решения; способность разрешать конфликты и смягчать разногласия; способность эффективно работать в качестве подчиненного; терпимость по отношению к различным стилям жизни окружающих; понимание плюралистической политики; готовность заниматься организационным и общественным планированием

Дж. Равен отмечает, что компетенции (мотивированные способности, по его определению) «могут заменять друг друга в качестве составляющих эффективного поведения» [79, с.253].

В отечественной психологии вопрос о компенсации одних способностей другими поднимался в работах Б. М. Теплова [94], В. Мерлина [50], К. К. Платонова [64], Е. П. Ильина [25], В. Д. Шадрикова [104], [105] и др. Резервы способностей человека позволяют компенсировать недостающие или недостаточно выраженные компоненты, необходимые для успешной деятельности, в том числе, учебной. Вместе с тем, определение компетенций и компетентности отличается от определения способностей как индивидуально-психологических особенностей человека, обеспечивающих успешность деятельности, обуславливающих лёгкость и быстроту приобретения знаний, умений и навыков, но не сводимых к знаниям, умениям и навыкам (Б. М. Теплов).

В отечественной педагогике и психологии понятию «компетенция» в наибольшей степени соответствует понятие «умение». К настоящему времени разработано множество концепций по проблеме формирования различных умений (И. И. Павлов, И. М. Сеченов, Л. С. Выготский, П. Я. Гальперин, А. Н. Леонтьев, Н. Ф. Талызина, Ю. К. Бабанский, Е. А. Милерян, С. Л. Рубинштейн, А. В. Усова, А. В. Запорожец, Д. Б. Эльконин). Нельзя не отметить, что понятие «умение» трактуется авторами по-разному. Как отмечает С. А. Тимаева [95], в отечественной науке понятие «умение» закрепилось за различными по своему масштабу изучаемыми явлениями. Одни исследователи связывают умение с поведением на уровне деятельности, другие – с выполнением действий, входящих в деятельность на правах самостоятельных единиц, а третьи – с овладением отдельными вспомогательными операциями в пределах некоторого действия. Понятию компетенции соответствует лишь первый вариант определения понятия «умение», в рамках которого умение – это готовность к выполнению какой-либо деятельности.

Так, Е. А. Милерян рассматривает умения как «сложные структурные образования личности, включающие чувственные, интеллектуальные, волевые, творческие, эмоциональные качества личности, обеспечивающие достижение поставленной цели деятельности в изменяющихся условиях ее протекания. В умениях навыки, как усвоенные действия, стали свойствами личности и ее способностями к новым действиям». Е. А. Милерян подчеркивает, что «умения представляют собой своеобразный синтез – структурный ансамбль природных и приобретенных свойств личности, поэтому каждому умению свойственна структура личностных качеств и свойств» [52]. Иными словами, «умение – это основанная на знаниях и навыках способность человека успешно достигать сознательно поставленной цели деятельности в изменяющихся условиях» [там же, с.36].

В концепции, разработанной К. К. Платоновым, умение, как «способность выполнять определенную деятельность и действие в новых условиях», является «конечной целью педагогического процесса, его завершением» [64]. Такое определение понятия «умение» созвучно с определением понятия «компетентность», которое в настоящее время используется в контексте результата образовательного процесса и готовности к профессиональной деятельности.

А. К. Маркова определяет следующие виды профессиональной компетентности [47, с.34]:

1. **Специальная компетентность** – владение собственно профессиональной деятельностью на достаточно высоком уровне, способность проектировать свое дальнейшее профессиональное развитие.

2. **Социальная компетентность** – владение совместной (групповой, кооперативной) профессиональной деятельностью, сотрудничеством, нормами профессионального общения; социальная ответственность за результат своего профессионального труда.

3. **Личностная компетентность** – владение приемами личностного самовыражения и саморазвития, средствами противостояния профессиональным деформациям личности.



**4. Индивидуальная компетентность** – владение приемами самореализации и развития индивидуальности в рамках профессии, готовность к профессиональному росту, способность к индивидуальному самосохранению, неподверженность профессиональному старению, умение рационально организовать свой труд, без перегрузок времени и сил, осуществлять труд ненапряженно, без усталости и даже с освежающим эффектом.

А. К. Маркова также отмечает, что существуют общие виды компетентности, необходимые для человека независимо от профессии (межпрофессиональные компоненты):

1. **Специальная компетентность** – способность к планированию производственных процессов, умение работать с компьютером, оргтехникой, чтение технической документации, ручные навыки.

2. **Личностная компетентность** – способность планировать свою труд деятельность, контролировать и регулировать ее, способность самостоятельно принимать решения, способность находить нестандартные решения (креативность), гибкое теоретическое и практическое мышление, умение видеть проблему, способность самостоятельно приобретать новые знания и умения.

3. **Индивидуальная компетентность** – мотивация достижения, ресурс успеха, стремление к качеству своей работы, способность к самомотивированию, уверенность в себе, оптимизм.

И. А. Зимняя определяет компетентность как личностную характеристику индивида, отражающую его способность использовать универсальные способы деятельности, основанную на совокупности научных знаний в конкретных жизненных ситуациях.

По определению И. А. Зимней компетентность – это основывающийся на знаниях, интеллектуально и личностно-обусловленный опыт социально-профессиональной жизнедеятельности человека [24]. Компетенции – это некоторые внутренние, потенциальные, сокрытые психологические новообразования: знания, представления, программы (алгоритмы) действий, систем цен-

ностей и отношений. И. А. Зимняя выделяет три группы основных компетенций.

1. Компетенции, относящиеся к самому человеку как личности, субъекту деятельности, общения:

- компетенции здоровьесбережения: знание и соблюдение норм здорового образа жизни, знание опасности курения, алкоголизма, наркомании, СПИДа; знание и соблюдение правил личной гигиены, обихода; физическая культура человека, свобода и ответственность выбора образа жизни;
- компетенции ценностно-смысловой ориентации в мире: ценности бытия, жизни; ценности культуры (живопись, литература, искусство, музыка); науки; производства; истории цивилизаций, собственной страны; религии;
- компетенции интеграции: структурирование знаний, ситуативно-адекватной актуализации знаний, расширения приращения накопленных знаний;
- компетенции гражданственности: знания и соблюдение прав и обязанностей гражданина; свобода и ответственность, уверенность в себе, собственное достоинство, гражданский долг; знание и гордость за символы государства (герб, флаг, гимн);
- компетенции самосовершенствования, саморегулирования, саморазвития, личностной и предметной рефлексии; смысл жизни; профессиональное развитие; языковое и речевое развитие; овладение культурой родного языка, владение иностранным языком.

2. Компетенции, относящиеся к социальному взаимодействию человека в обществе:

- компетенции социального взаимодействия: с обществом, общностью, коллективом, семьей, друзьями, партнерами, конфликты и их погашение, сотрудничество, толерантность, уважение и принятие Другого (раса, национальность, религия, статус, роль, пол), социальная мобильность;
- компетенции в общении: устном, письменном, диалог, монолог, порождение и восприятие текста; знание и соблюдение тради-

ций, ритуала, этикета; кросскультурное общение; деловая переписка; делопроизводство, бизнес-язык; иноязычное общение, коммуникативные задачи, уровни воздействия на реципиента.

### 3. Компетенции, относящиеся к деятельности человека:

- компетенции познавательной деятельности: постановка и решение познавательных задач; нестандартные решения, проблемные ситуации – их создание и разрешение; продуктивное и репродуктивное познание, исследование, интеллектуальная деятельность;
- компетенции деятельности: игра, учение, труд; средства и способы деятельности: планирование, проектирование, моделирование, прогнозирование, исследовательская деятельность, ориентация в разных видах деятельности;
- компетенции информационных технологий: прием, переработка, выдача информации; преобразование информации (чтение, конспектирование), массмедийные, мультимедийные технологии, компьютерная грамотность; владение электронной, Интернет-технологией.

Каково же соотношение понятий «компетенция» и «компетентность»?

И. А. Зимняя определяет компетенции как «актуальные компетентности», т.е. компетенции суть проявления компетентности в различных аспектах:

- мотивационно-ценностном (готовность и отношение к цели проявления компетентности);
- когнитивном (владение необходимыми знаниями);
- деятельностном (опыт проявления компетентности в учебной и профессиональной деятельности);
- эмоционально-волевом (эмоционально-волевая регуляция процесса и результата проявления компетентности).

В. Е. Медведев и Ю. Г. Татур объединяют понятия компетенции и частной компетентности. В соответствии с их определением [49], компетенция (частная компетентность) – это обобщенная характеристика личности специалиста, определяющая проявленную им

готовность использовать свой потенциал (знания, умения, опыт и личностные качества) для успешной деятельности в определенной социальной или профессиональной области.

В общем смысле термин «компетентность» употребляется тогда, когда речь идет о **единой интегральной характеристике специалиста**. Например, компетентный юрист, компетентный врач, компетентный преподаватель. В тех же случаях, когда надо охарактеризовать компетентность специалиста в одной из областей его деятельности, используется термин «компетенция». Так, инженер может продемонстрировать свою компетенцию в области конструирования, организации производства, эксплуатации объектов и т.п. В компетентности профессионала интегрируются компетенции (частные компетентности), при этом недостающие или недостаточно выраженные компоненты, необходимые для достижения цели, компенсируются за счёт более развитых компонентов компетентности. Одновременно происходит формирование, развитие и совершенствование компетенций в процессе приобретения человеком личного опыта соответствующего вида деятельности.

Таким образом, **компетентность** представляет собой интегральную характеристику специалиста, в структуру которой входят:

- когнитивный компонент (знания);
- функциональный компонент (умения, навыки);
- мотивационно-ценностный компонент (отношение к осуществляемой деятельности);
- личностно-обусловленный опыт социально-профессиональной жизнедеятельности человека.

Для определения **компетенции** можно использовать два подхода. Если рассматривать компетенцию как характеристику специалиста, в соответствии с определением Дж. Равена: компетенция – это «мотивированная способность». Если же рассматривать компетенцию с позиции профессиональной деятельности, то компетенция представляет собой проявление общей компетентности специалиста в одной из областей его деятельности (Ю. Г. Татур).

Анализ содержания профессиональных и общекультурных компетенций Федерального Государственного образовательного стандарта высшего профессионального образования показывает, что любую компетенцию из их числа можно отнести к одному из компонентов **личностной компетентности**, которая определяется, как **интегральная характеристика личности обучающегося** и имеет следующую структуру:

**индивидуально-психологический компонент** (личностные особенности, интеллектуальные и творческие способности, лежащие в основе готовности и способности обучающегося к обучению и саморазвитию);

**мотивационно-ценностный компонент** (сформированность мотивации к обучению и познанию, ценностно-смысловые установки);

**деятельностный компонент** (учебные умения, навыки, сформированность универсальных учебных действий);

**коммуникативный компонент** (коммуникативные способности, навыки межличностного взаимодействия в образовательном процессе).

Формирование личностной компетентности обучающихся осуществляется на допрофессиональном и профессиональном этапах непрерывного образования.

Сформированность указанных четырёх компонентов личностной компетентности является необходимым условием адаптации студентов к обучению в вузе и обеспечения непрерывности образования. Степень и особенности адаптации (дезадаптации) определяются уровнем сформированности всех компонентов личностной компетентности. Деадаптация (полная или частичная) приводит к снижению эффективности обучения вплоть до отчисления. Полная деадаптация может привести к психическим и соматическим расстройствам. Для полноценного формирования компетентности студентов исследовательского университета необходимо не только совершенствовать материально-техническую базу для научной, учебной, внеучебной и общественной деятельности студентов, но

также широко использовать в учебно-воспитательном процессе инновационные методы обучения и средства психологического сопровождения образовательного процесса.

По результатам оценки сформированности компонентов личностной компетентности разрабатываются коррекционные и развивающие программы. Уровни сформированности отдельных структурных компонентов определяют структуру личностной компетентности каждого студента и являются основой для дифференциации и индивидуализации программы психологического сопровождения.

**Таблица 5.2. Соответствие профессиональных и личностных компетенций структурным компонентам личностной компетентности студентов**

Структурные компоненты личностной компетентности	Профессиональные компетенции	Общекультурные компетенции
индивидуально-психологический		
мотивационно-ценностный		
деятельностный		
коммуникативный		

### **Контрольные вопросы и задания**

1. Дайте определения понятиям «компетентность» и «компетенция». Приведите примеры профессиональных и общекультурных компетенций выпускника исследовательского университета.
2. Какие составляющие образовательной среды исследовательского университета создают условия для формирования профессиональных компетенций?



3. *Какие составляющие образовательной среды исследовательского университета создают условия для формирования общекультурных компетенций?*
4. *Выберите любую дисциплину из учебного плана. Какие профессиональные и общекультурные компетенции формируются у студентов в ходе изучения этой дисциплины? Какие методы обучения следует использовать для повышения эффективности формирования этих компетенций?*
5. *Заполните таблицу 5.2, сгруппировав общекультурные компетенции ФГОС по своему направлению подготовки в соответствии с компонентами личностной компетентности.*

## **5.2. Критерии и показатели сформированности общекультурных и профессиональных компетенций студентов**

Оценка сформированности личностной компетентности осуществляется с помощью психодиагностического метода (методики подбираются в соответствии с оцениваемыми компонентами и возрастом обучающихся), а также методов экспертных оценок, анкетного опроса, наблюдения. Критериями сформированности компонентов личностной компетентности являются соответствие показателей сформированности личностной компетентности 1) возрастным психологическим нормам, 2) групповым нормам, 3) требованиям учебной программы, 4) дисциплинарным нормам.

Оценка сформированности личностной компетентности студента может производиться

- методом экспертных оценок преподавателями и кураторами;
- с помощью психодиагностических методик и метода анкетного опроса психологической службой вуза;
- обучающимися самостоятельно.

**Индивидуально-психологический компонент** личностной компетентности, показателем сформированности которого являются профессионально важные качества личности, оценивается с помощью психодиагностических методик, экспертных оценок, в ходе деловых игр.

**Мотивационно-ценностный компонент**, показателем сформированности которого является учебно-профессиональная мотивация и система ценностей студента, оценивается с помощью психодиагностических методик и метода анкетного опроса.

**Деятельностный компонент**, показателем сформированности которого является самостоятельность и исследовательский характер учебной деятельности, оценивается с помощью методов экспертных оценок, анкетного опроса.

**Коммуникативный компонент**, показателем сформированности которого является умение налаживать продуктивные интерперсональные взаимоотношения в учебной, общественной и практической деятельности, оценивается с помощью методов социометрии, экспертных оценок, анкетного опроса.

При оценке сформированности компетенций важную роль играет активная позиция самого студента, заинтересованного в качестве получаемого им образования. **В данном контексте существует принципиальное отличие оценки сформированности компетенций от традиционной школьной оценки, как итоговой характеристики уровня знаний.** Результаты психодиагностики, экспертной оценки приобретают развивающее значение для студента, стремящегося к личностному росту. Важнейшее значение приобретает также самоконтроль и самооценка обучающимися своих компетенций. Студент, заинтересованный в получении качественного профессионального образования, является субъектом образовательной деятельности, подбирая с помощью консультанта (преподавателя, куратора, сотрудника психологической службы) те виды учебной, научной и практической деятельности, которые способствуют формированию соответствующих компетенций.

Показатели сформированности компонентов личностной компетентности студентов и методы их оценки представлены в табл. 5.3.

Рассмотрим примеры методов оценки отдельных показателей сформированности личностной компетентности.

1. Профессионально значимые личностные качества: 16ФЛО (Р. Кеттелл); Многоуровневый личностный опросник «Адаптив-

ность» (А.Г. Маклаков, С.В. Чермянин), шкала «Нервно-психическая устойчивость»; Многомерный тест смысловых ориентаций (Д. А. Леонтьев); Методика исследования совладающего поведения (Р. Лазарус).

2. Коммуникативные способности: многоуровневый личностный опросник «Адаптивность» (шкала «Коммуникативные способности»; методика диагностики интерперсональных отношений Т.Лири; методика Е.Залюбовской; методика КОС Б.Федоришина, социометрия.

**Таблица 5.3 Оценка сформированности личностной компетентности студентов**

Структурные компоненты компетентности студента	Показатели сформированности	Методы оценки
индивидуально-психологические особенности (когнитивные, личностные)	знания	анализ успеваемости
	профессионально значимые личностные качества	психодиагностический метод, деловые игры, экспертные оценки
мотивационно-ценностный	учебно-профессиональная мотивация, ценности	психодиагностический метод
деятельностный	самостоятельность и научный характер учебной деятельности	анкетный опрос, психологическая оценка
	участие в НОУ, общественной деятельности, разработка научно-исследовательских и социальных проектов	анализ результатов научно-исследовательской и общественной деятельности
коммуникативный	коммуникативные способности	психодиагностический метод, деловые игры

В качестве примера оценки коммуникативной компетентности студентов приведём результаты обследования социально активных студентов, участвующих в студенческом самоуправлении вуза. Анализ коммуникативной компетентности студентов показал, что в выборке социально активных студентов подавляющее большинство респондентов показали высокий и очень высокий уровень сформированности коммуникативных и организаторских способностей

(76% и 88% соответственно) по опроснику КОС (коммуникативные и организаторские способности) Б. Федоришина. Это свидетельствует о том, что социально активные студенты быстро ориентируются в трудных ситуациях, свободно ведут себя в новом коллективе, инициативны, предпочитают в важном деле или в сложной ситуации принимать самостоятельные решения, отстаивают свое мнение и добиваются, чтобы оно было принято товарищами, настойчивы в деятельности, которая их привлекает. Заметим, что в контрольной группе студентов-психологов 3 курса, для которых коммуникативная компетентность является профессионально важным личностным качеством и целенаправленно развивается в процессе обучения, высокий уровень коммуникативных способностей показали 59% респондентов (различия статистически незначимы), а высокий уровень организаторских способностей – лишь 33% (различия статистически значимы,  $p < 0,001$ ). Полученные результаты отражают важность самостоятельной социально активной деятельности для формирования организаторских способностей и таких компетенций, как ОК1б (умение работать самостоятельно и в коллективе, руководить людьми и подчинять личные интересы общей цели); ОК8м (способность быстро адаптироваться к любым ситуациям); ПК13м (способность к управлению и руководству научной работой коллективов). Сформированность лидерских компетенций (ОК7м) у социально активных студентов была также подтверждена с помощью методики диагностики интерперсональных отношений Т. Лири. У социально активных студентов отмечена также большая выраженность коллективистской направленности личности (методика Е. Залюбовской).

Решение вопроса об психологической оценке сформированности у студентов общекультурных и профессиональных компетенций, указанных в федеральном государственном образовательном стандарте невозможно без активной позиции обучающихся, самоконтроля и самооценки ими своей учебной деятельности и её результатов в виде компетенций.

В заключение отметим, что многие из профессиональных и общекультурных компетенций, перечисленных в федеральном государственном образовательном стандарте, могут быть сформированы лишь в ходе практик. Вместе с тем, практика – это один из наименее стандартизированных видов учебной деятельности, содержание которого для студентов даже одной академической группы может существенно различаться. Поэтому на протяжении всего периода обучения в вузе, у каждого студента должен быть единый дневник для всех практик. В этом дневнике студент самостоятельно, под руководством преподавателей и руководителей практики от предприятия, оценивает сформированность у себя профессиональных и общекультурных компетенций, указанных в федеральном образовательном стандарте, и определяет те компетенции, над развитием которых нужно работать в дальнейшей учебной и практической деятельности.

### **Контрольные вопросы и задания**

- 1. Предложите методы оценки сформированности деятельностного компонента личностной компетентности студентов.*
- 2. С помощью каких психодиагностических методик можно оценить сформированность мотивационно-ценностного компонента личностной компетентности? Предложите методику оценки каждой компетенции, относящейся к мотивационно-ценностному компоненту (табл. 5.2).*
- 3. Какие факторы 16ФЛО позволяют оценить профессионально важные личностные качества студента (выпускника) вуза по Вашему направлению подготовки?*

## **Глава 6.**

### **Методы первичной обработки данных психологического обследования**

#### **6.1. Первичная обработка данных психологического обследования**

Начальный этап обработки данных заключается в визуальном и описательном анализе значений показателей применяемой психодиагностической методики. Для этого данные табулируются, упорядочиваются, производится их группировка, вычисляются простые описательные статистики, строятся графики.

#### **Табулирование данных**

До анализа и интерпретации количественных данных их необходимо обобщить. В качестве примера рассмотрим результаты диагностики когнитивной сферы группы студентов с помощью теста «Стандартные прогрессивные матрицы» Дж.Равена. Данные сведены в таблицу и упорядочены в алфавитном порядке (табл.6.1).

Первый этап обработки данных – это обычно упорядочивание оценок от максимальной до минимальной. Такое представление называют несгруппированным рядом. В результате получается таблица, иллюстрирующая порядок испытуемых по степени наличия у них измеряемого психического качества.

При большом объеме выборки имеет смысл разбить всю группу обследуемых на 3-5 частей и в индивидуальных результатах указать, в какую ранговую группу попал каждый испытуемый (табл. 6.2). Список можно сократить, объединяя в одну группу испытуемых с одинаковой оценкой. В этом случае получается таблица распределения частот (или вариационный ряд значений признака), в



которой каждому значению признака соответствует частота его встречаемости в данной выборке (табл. 6.3).

**Таблица 6.1 Первичные данные обследования.**

Фамилия, имя	Оценка
А. Евгений	53
А. Илья	46
А. Ирина	45
А. Николай	51
Г. Денис	45
Е. Анна	51
И. Юлия	53
К. Андрей	56
К. Ирина	50
К. Роман	44
М. Василий	51
Н. Юрий	53
П. Николай	54
Т. Михаил	42
Ч. Юлия	48

**Таблица 6.2. Упорядоченное множество данных.**

Фамилия, имя	Оценка	Группа
К. Андрей	56	1
П. Николай	54	
И. Юлия	53	
Н. Юрий	53	
А. Евгений	53	
А. Николай	51	2
Е. Анна	51	
М. Василий	51	
К. Ирина	50	
Ч. Юлия	48	
А. Илья	46	3
А. Ирина	45	
Г. Денис	45	
К. Роман	44	
Т. Михаил	42	

**Таблица 6.3. Распределение частот**

Сумма	Частота
56	1
54	1
53	3
51	3
50	1
48	1
46	1
45	2
44	1
42	1

Для большого числа оценок (100 или более) на следующем этапе может иметь смысл обобщение данных. Как правило, широкий диапазон оценок целесообразно сгруппировать их по величинам, например, в группы, объединяющие все оценки от 56 до 60 включительно, от 51 до 55 и т.д. Каждая такая группа называется *разрядом оценок*. В случае полного размещения по группам обычно говорят о *распределении сгруппированных частот*. Четкого правила выбора количества разрядов не существует, но рекомендуется образовывать не менее 12 и не более 15 разрядов. В нашем случае, вследствие малого объема выборки и недостаточно широкого диапазона оценок, нет необходимости группировать данные, но в табл.6.4 это сделано в качестве иллюстрации.

В зависимости от целей исследования выборка испытуемых может быть разбита на две и более подгрупп (например, нам требуется изучить уровень развития когнитивной сферы студентов мужского и женского пола либо сопоставить между собой группы студентов различных специализаций). Определим высокий и низкий (для данной выборки!) уровни результативности выполнения заданий теста в соответствии с числом набранных баллов (табл.6.4) и посмотрим, каким образом этот уровень соотносится с полом испытуемых (табл.6.5). Для наглядного представления такого рода данных используются таблицы сопряженности (табл.6.6).

**Таблица 6.4. Распределение сгруппированных частот**

Диапазон оценок	Частота	Уровень
56–60	1	В
51–55	7	
46–50	3	Н
41–45	4	

**Таблица 6.5. Данные, подготовленные к построению таблицы сопряженности**

Фамилия, имя	Пол	Оценка	Уровень
А. Евгений	м	53	В
А. Илья	м	46	Н
А. Ирина	ж	45	Н
А. Николай	м	51	В
Г. Денис	м	45	Н
Е. Анна	ж	51	В
И. Юлия	ж	53	В
К. Андрей	м	56	В
К. Ирина	ж	50	Н
К. Роман	м	44	Н
М. Василий	м	51	В
Н. Юрий	м	53	В
П. Николай	м	54	В
Т. Михаил	м	42	Н
Ч. Юлия	ж	48	Н

**Таблица 6.6. Таблица сопряженности**

Уровень развития когнитивной сферы	Пол	
	М	Ж
Высокий	6 (60%)	2 (40%)
Низкий	4 (40%)	3 (60%)

Таблицы сопряженности имеют не только иллюстративное, но и аналитическое значение, позволяя обнаруживать различные аспекты связей между переменными. Так, таблица 6.6 (в отличие табл. 6.5) позволяет заметить интересный факт: 60% испытуемых мужского пола получили высокую оценку по тесту Равена, а 40% – низкую; для девушек соотношение обратное: 40% – высокая оценка, 60% – низкая. Говорят ли эти данные о том, что в данной выборке испытуемых уровень развития когнитивной сферы девушек в целом ниже, чем у юношей? Для доказательства статистической достоверности этих различий визуального анализа таблицы сопряженности и арифметического сравнения процентных долей недостаточно, требуется применять статистический аппарат. Решение такого рода задач будет рассмотрено в главе 7.

В некоторых случаях для визуального анализа больших массивов данных используют сложные таблицы сопряженности, в которых данные сгруппированы по двум ранжированным признакам. Примером такой таблицы (табл.6.7) может служить представление классических данных Ф.Гальтона [22, с. 38]. Отметим, что анализировать представленную таблицу достаточно сложно, если принять во внимание разное общее количество испытуемых в ростовых категориях.

**Таблица 6.7. Соотношение роста детей и родителей (по данным Ф. Гальтона)**

Показатели		Рост детей								Всего
		154	159	164	169	174	180	185	190	
Рост родителей	188							4		4
	183			1	7	11	17	20	6	62
	178	1	2	21	48	83	66	22	8	251
	173	1	15	56	130	148	69	11		430
	168	1	15	19	56	41	11	1		144
	163	2	7	10	14	4				37
Всего		5	39	107	255	287	163	58	14	928

Очевидно, большей информативностью будет обладать таблица относительных частот встречаемости детей различного роста у ка-

ждой категории родителей (табл.6.8), либо таблица относительных частот роста родителей у каждой категории детей. Различные способы группировки данных и анализ соответствующих таблиц сопряженности помогают исследователю сделать предположения о различных (и зачастую неожиданных) связях, тенденциях. При дальнейшей обработке полученных данных истинность этих предположений проверяется с помощью статистических методов.

**Таблица 6.8. Относительная частота встречаемости детей различного роста у каждой категории родителей**

Показатели		Рост детей								Всего
		154	159	164	169	174	180	185	190	
Рост родителей	188	0%	0%	0%	0%	0%	0%	<b>100%</b>	0%	100%
	183	0%	0%	2%	11%	18%	27%	<b>32%</b>	10%	100%
	178	0%	1%	8%	19%	<b>33%</b>	26%	9%	3%	100%
	173	0%	3%	13%	30%	<b>34%</b>	16%	3%	0%	100%
	168	1%	10%	13%	<b>39%</b>	28%	8%	1%	0%	100%
	163	5%	19%	27%	<b>38%</b>	11%	0%	0%	0%	100%
	Всего	1%	4%	12%	27%	31%	18%	6%	2%	100%

### Ранжирование данных

При выполнении статистической обработки данных зачастую требуется произвести их ранжирование. Суть процедуры ранжирования заключается в проставлении номеров (рангов) значениям в упорядоченном множестве данных. Отметим, что поскольку ранжирование является формальной процедурой, то в зависимости от целей, в которых производится ранжирование, самый высокий ранг можно присваивать как наибольшему, так и наименьшему значению (соответственно, ранжирование будет производиться «по возрастанию» или «по убыванию»). Рассмотрим тот вариант процедуры ранжирования, когда самый высокий ранг присваивается наибольшему значению, поскольку именно этот вариант ранжирования

понадобится нам при вычислении непараметрических статистических критериев.

**Правила ранжирования:**

1. Упорядочить данные по возрастанию и пронумеровать их.
2. Ранг равен порядковому номеру значения, если это значение встречается только 1 раз.
3. Если значение встречается более 1 раза, то ранги этих значений будут одинаковыми и равными среднему арифметическому значению их порядковых номеров.
4. Сумма полученных рангов равна  $N*(N+1)/2$ , где  $N$  – объём выборки (проверка правильности ранжирования).

В таблице 6.9 представлен результат ранжирования представленных в табл. 6.1 данных.

**Таблица 6.9. Ранжирование результатов выполнения студентами теста Равена**

№	Фамилия, имя	Оценка	Количество одинаковых значений	Вычисление ранга	Ранг
1	Т. Михаил	42	1	1	1
2	К. Роман	44	1	2	2
3	А. Ирина	45	2	$(3+4)/2=3,5$	3,5
4	Г. Денис	45			3,5
5	А. Илья	46	1	5	5
6	Ч. Юлия	48	1	6	6
7	К. Ирина	50	1	7	7
8	А. Николай	51	3	$(8+9+10)/3=9$	9
9	Е. Анна	51			9
10	М. Василий	51			9
11	И. Юлия	53	3	$(11+12+13)/3=12$	12
12	Н. Юрий	53			12
13	А. Евгений	53			12
14	П. Николай	54	1	14	14
15	К. Андрей	56	1	15	15
Сумма рангов					120

Проверка правильности ранжирования осуществляется по формуле:

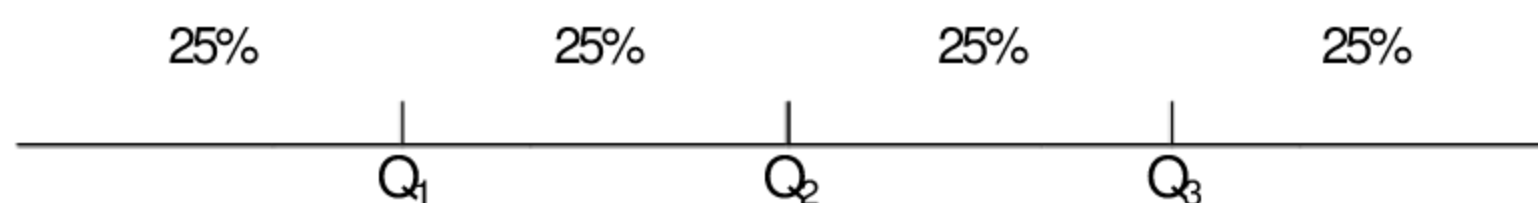


$$N*(N+1)/2=15*16/2=120 \quad (6.1)$$

### Квантили

Несгруппированный ряд значений переменной может быть описан с помощью квантилей. Квантиль – это точка на числовой шкале, которая делит совокупность наблюдений на две группы с известными пропорциями в каждой из них. Квантиль – это общее понятие, конкретизация которого зависит от того, на сколько равных частей разделено упорядоченное множество значений. При описании результатов психологических обследований с помощью квантилей, как правило, используются квартили (шкала делится на 4 равные части) и процентиля (шкала делится на 100 равных частей).

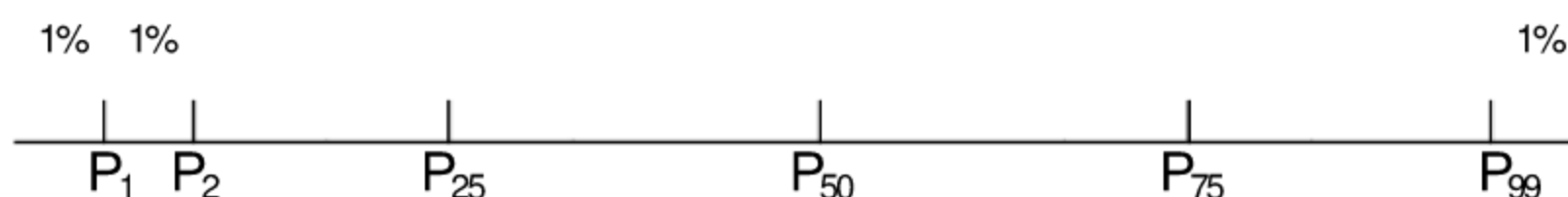
Существуют три квартиля ( $Q_1$ ,  $Q_2$ ,  $Q_3$ ). Они делят группу наблюдений на четыре равные части (кварти): четвертая часть наблюдений лежит ниже  $Q_1$ , половина наблюдений лежит ниже  $Q_2$ , а три четверти наблюдений – ниже  $Q_3$  (рис. 6.1). Выражение  $Q_1=7$  означает, что по результатам проведенного обследования четверть испытуемых набрали 7 баллов и менее.



**Рис. 6.1. Шкала квартилей**

Квартили очень удобны для описания множества наблюдений. Так, значение  $Q_2$  определяет середину несгруппированного ряда, а с помощью  $Q_1$  и  $Q_3$  вычисляется квартильный размах, используемый в некоторых случаях для характеристики вариативности признака.

Кроме квартилей, для анализа и представления данных психологических исследований используются процентиля. 99 возможных процентилей ( $P_1, \dots, P_{99}$ ) делят множество наблюдений на 100 частей с равным числом наблюдений в каждой (рис. 6.2).



**Рис. 6.2. Шкала процентилей**

Выражение  $P_{15} = 20$  означает, что по результатам обследования оценка 15% испытуемых составляет 20 баллов и ниже. Между квантилями и процентилями существует очевидная связь: так, если 25% всех наблюдений находится ниже  $P_{25}$ , и то же справедливо для  $Q_1$ , то  $Q_1 = P_{26}$ . Аналогично,  $Q_2 = P_{50}$  и  $Q_3 = P_{76}$ .

Процентили часто используются для представления тестовых норм. В качестве примера приведем данные стандартизации теста Равена на выборке взрослых жителей Великобритании (табл.6.10). Из таблицы видно, что 75% выборки набрало не более 57 баллов и всего лишь 10% – меньше 44.

**Таблица 6.10. Нормы теста Д. Равена в процентилях**

Процентиль	5	10	25	50	75	90	95
Число баллов	39	44	49	54	57	58	59

Процентильные нормы, представленные в таблице 6.10, позволяют оценить успешность выполнения теста Равена каждым испытуемым. Как известно, тест «Стандартные прогрессивные матрицы Равена» состоит из 60 заданий. И результат 54 балла из 60, который на первый взгляд кажется высоким, на самом деле является весьма посредственным, поскольку 50% испытуемых выборки стандартизации показали более высокий результат.

### **Контрольные вопросы и задания**

1. *Определите шкалу измерения:*
  - время (в секундах) выполнения теста Д.Равена;
  - образование испытуемых (среднее, бакалавр, магистр);
  - IQ в тесте Д.Векслера;
  - тип темперамента в опроснике Г.Айзенка (EPQ);
  - число верно выполненных заданий в тесте Д.Равена;
  - число верно выполненных заданий в тесте Д.Равена, как показатель интеллектуального развития;

- пол испытуемых;
- возраст испытуемых;
- ФИО испытуемых;
- умственный возраст в тесте А.Бине.

2. Проранжируйте данные, представленные в таблице 6.11.

**Таблица 6.11 Частота встречаемости различных оценок в группе.**

Шифр испытуемого	Балл по тесту	Ранг
А	25	
Б	13	
В	24	
Г	18	
Д	20	
Е	15	
Ж	13	
З	21	
И	22	
К	11	
Л	13	
М	19	
Н	24	
О	17	
П	13	

3. Что означают выражения:  $P_{85} = 85$ ;  $Q_1 = 15$ ?

## 6.2. Графическое представление данных

Результаты психодиагностического обследования можно проиллюстрировать с помощью различных типов графиков и диаграмм. Графическая иллюстрация результатов – необходимое дополнение к таблицам, т.к. многие закономерности, не видимые в численном виде, отчетливо проявляются на графиках. Грамотный выбор нужного типа графика очень важен, он зачастую определяет «лицо» проведенного исследования, ставит необходимые акценты. К сожалению, многие начинающие исследователи не придают должного значения выбору наиболее подходящей иллюстрации своих результатов. В результате получаются картинки, которые не несут ника-

кой смысловой информации, представляя собой, например, громоздкий частокор всех оценок каждого испытуемого либо «усредненный личностный профиль всех сотрудников организации» (список примеров можно продолжить).

Существует множество способов графического представления данных – это двумерные и трехмерные варианты разнообразных гистограмм, диаграмм размахов и диапазонов, столбчатых и круговых диаграмм, линейные и ленточные графики, диаграммы рассеяния и т.д. Далее будут рассмотрены двумерные способы графического представления данных, которые чаще всего используются для иллюстрации результатов психодиагностических обследований. Главное правило грамотного построения графика – он должен соответствовать результатам аналитической работы исследователя.

### **Гистограмма**

Гистограмма (термин введен К.Пирсоном в 1895 г.) является одним из наиболее любимых психологами и социологами типом графического представления данных. Гистограммы используются, в первую очередь, для большого числа наблюдений (более 100 испытуемых) и иллюстрируют распределение значений показателей психодиагностического обследования по интервалам группировки (т.е. как часто исследуемые переменные принимают значения из различных интервалов). Гистограмма наглядно показывает, какие значения (диапазоны значений) переменная принимает наиболее часто, где сконцентрировано большинство наблюдений, является ли распределение симметричным и т.д. Различают простые и составные гистограммы.

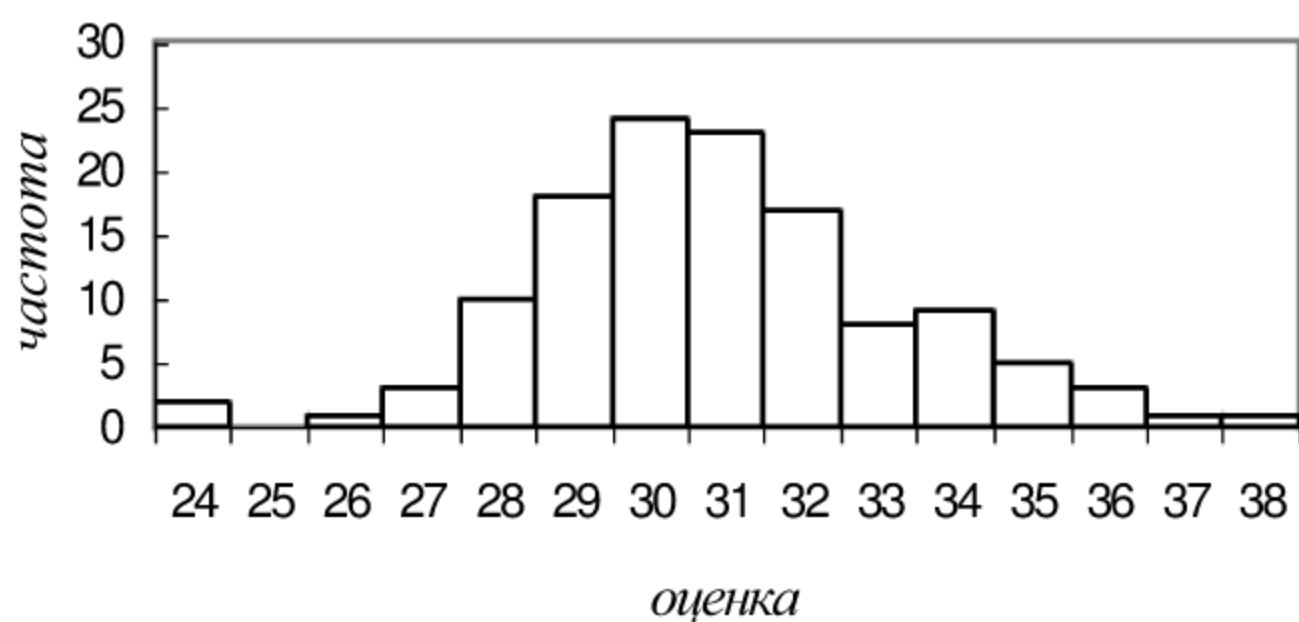
Для построения простой гистограммы на оси абсцисс откладываются оценки (или интервалы оценок), полученные испытуемыми в ходе психодиагностического обследования. Над каждой оценкой строится столбик, высота которого определяется частотой встречаемости этой оценки. В результате получается графическое изображение частотного распределения.

Рассмотрим пример. Преподаватель предложил 125 студентам контрольное задание, состоящее из 40 вопросов. В качестве оценки выбиралось количество вопросов, на которые были получены правильные ответы. Первичная обработка данных заключается в упорядочивании студентов в соответствии с полученными оценками и подсчете частоты встречаемости каждой оценки (табл.6.11).

Для построения гистограммы распределения частот по оси абсцисс отложим оценки студентов. Минимальная оценка равна 24, максимальная – 38. Оценку 24 получили двое студентов (частота  $f=2$ ), поэтому высота столбика над 24 будет равна 2. Оценку 25 не получил никто ( $f=0$ ), следовательно, столбика над 25 не будет. Оценку 30 получили 24 студента ( $f=24$ ), и высота столбика над 30 будет равна 24. В результате этих построений получается гистограмму распределения частот (рис.6.3).

**Таблица 6.11 Частота встречаемости различных оценок в группе.**

Оценка	Частота $f$
38	1
37	1
36	3
35	5
34	9
33	8
32	17
31	23
30	24
29	18
28	10
27	3
26	1
25	0
24	2



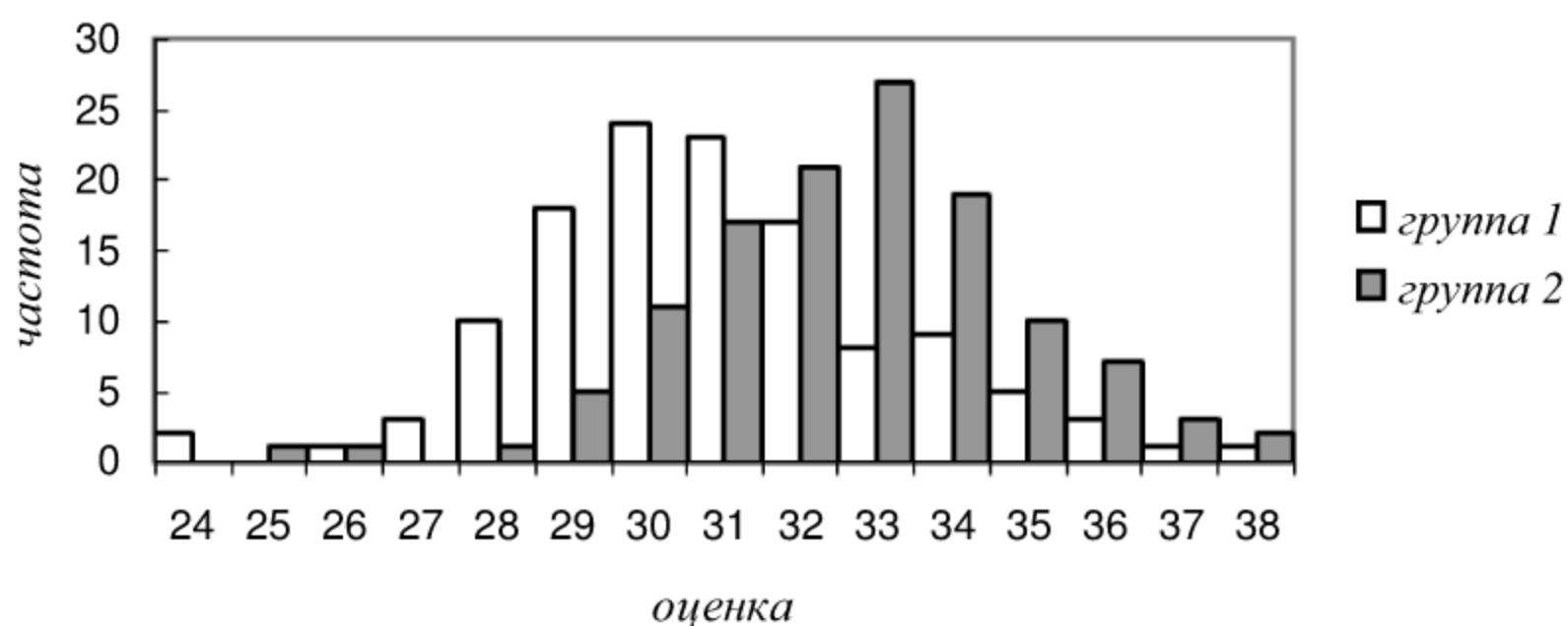
**Рис. 6.3. Простая гистограмма распределения частот**

Рассмотрим еще один пример. Контрольное задание, о котором шла речь в предыдущем примере, было предложено другой группе студентов. В табл. 6.12 представлены оценки студентов из обеих групп, а на рис. 6.4 – составная гистограмма распределения частот оценок в каждой из групп, в которой каждой оценке соответствует два столбика высотой, равной частоте оценки в каждой группе.

**Таблица 6.12 Распределение оценок в двух группах.**

Группа 1		Группа 2	
Оценка	Частота $f_1$	Оценка	Частота $f_2$
38	1	38	2
37	1	37	3
36	3	36	7
35	5	35	10
34	9	34	19
33	8	33	27
32	17	32	21
31	23	31	17
30	24	30	11
29	18	29	5
28	10	28	1
27	3	27	0
26	1	26	1
25	0	25	1
24	2	24	0





**Рис. 6.4 Составная гистограмма распределения частот**

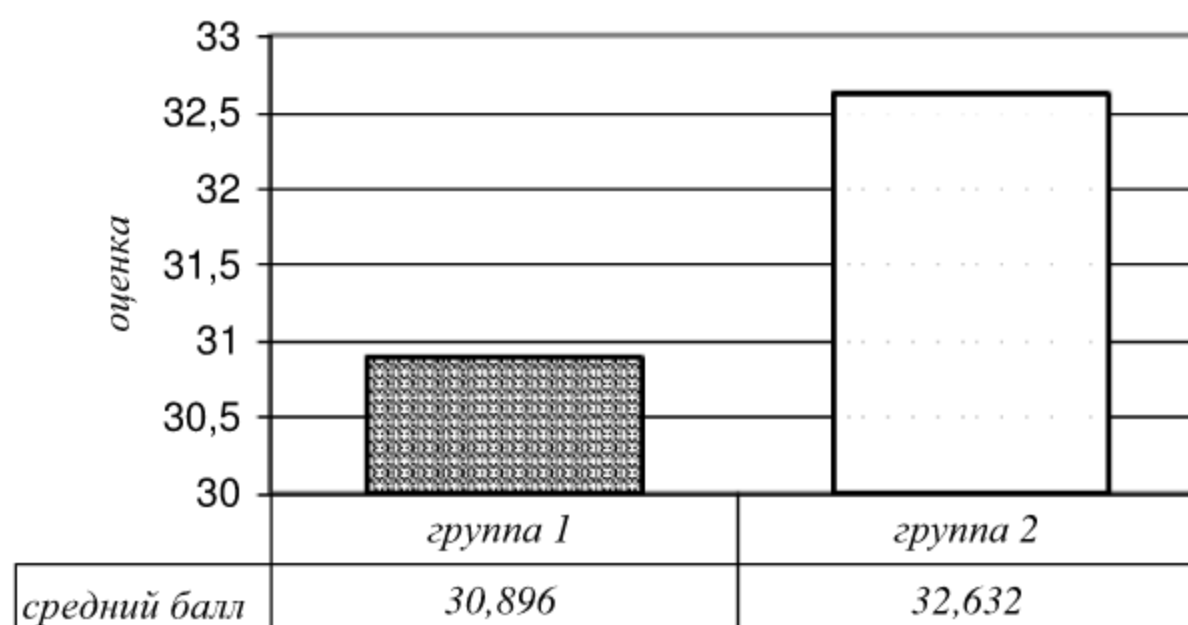
Составные гистограммы облегчают визуальное сопоставление распределения значений в обследованных выборках.

### Столбиковая диаграмма

Столбиковая (столбчатая) диаграмма используется для иллюстрации различий в групповых показателях исследуемого признака. При ее построении высота каждого столбика определяется средним значением признака в каждой группе.

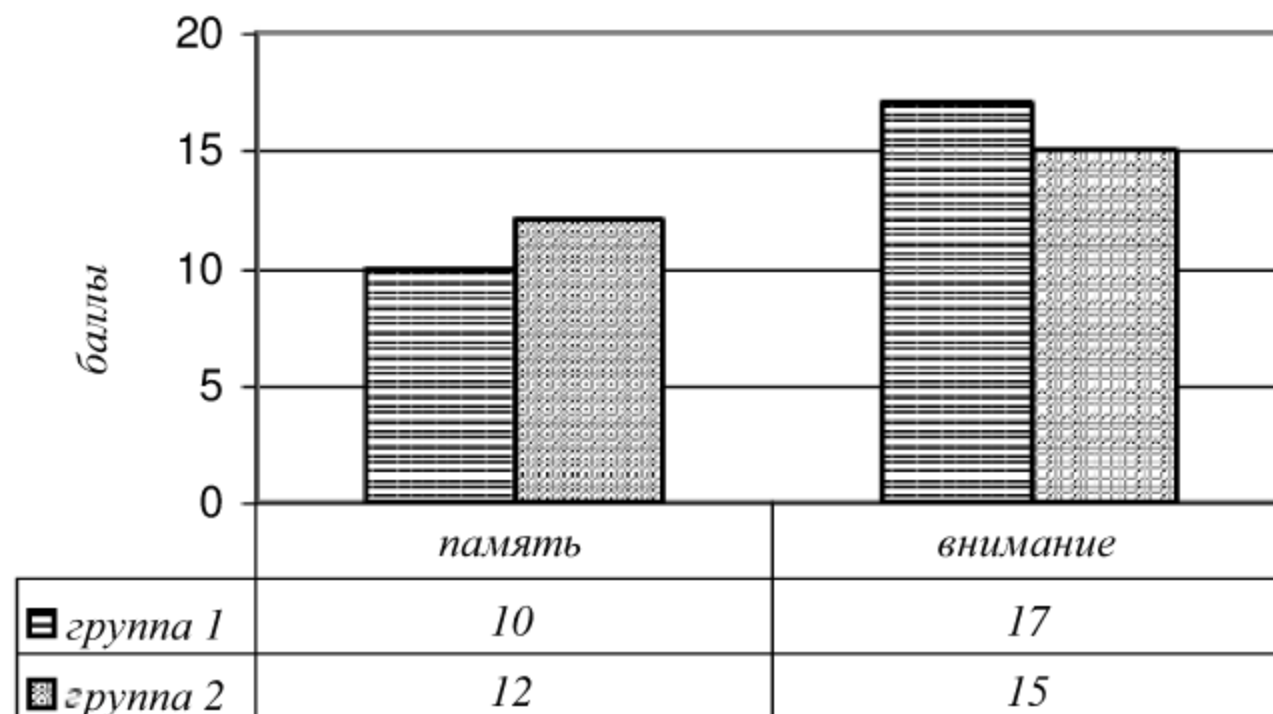
Продолжим рассмотрение примера с оценками студентов. По данным, представленным в табл. 6.11, вычислим среднегрупповые оценки. Для этого в каждой группе умножим каждую оценку на частоту ее встречаемости, сложим полученные числа и разделим сумму на число испытуемых. Средняя оценка в группе 1 равна 30,9, а в группе 2 – 31,6. Диаграмма, иллюстрирующая различие среднегрупповых оценок, представляет собой два столбика высотой, равной средней оценке по каждой группе (рис.6.5). Внизу столбиковой диаграммы можно расположить таблицу средних оценок по каждой группе.

Отметим, что подобная столбиковая диаграмма – это всего лишь **иллюстрация** различий оценок в группах испытуемых. Доказательством **достоверности** этих различий графическое представление данных **не является**. Такого рода вывод можно сделать лишь по результатам статистической обработки имеющихся числовых данных.



**Рис. 6.5 Результативность выполнения контрольного задания студентами групп 1 и 2**

Аналогичным образом строятся столбиковые диаграммы, иллюстрирующие различия в группах испытуемых по нескольким показателям. Предположим, что экспериментатор обследовал уровень развития памяти и внимания в двух группах испытуемых. Столбиковая диаграмма, иллюстрирующая соотношение уровней памяти и внимания в этих группах, представлена на рис. 6.6.



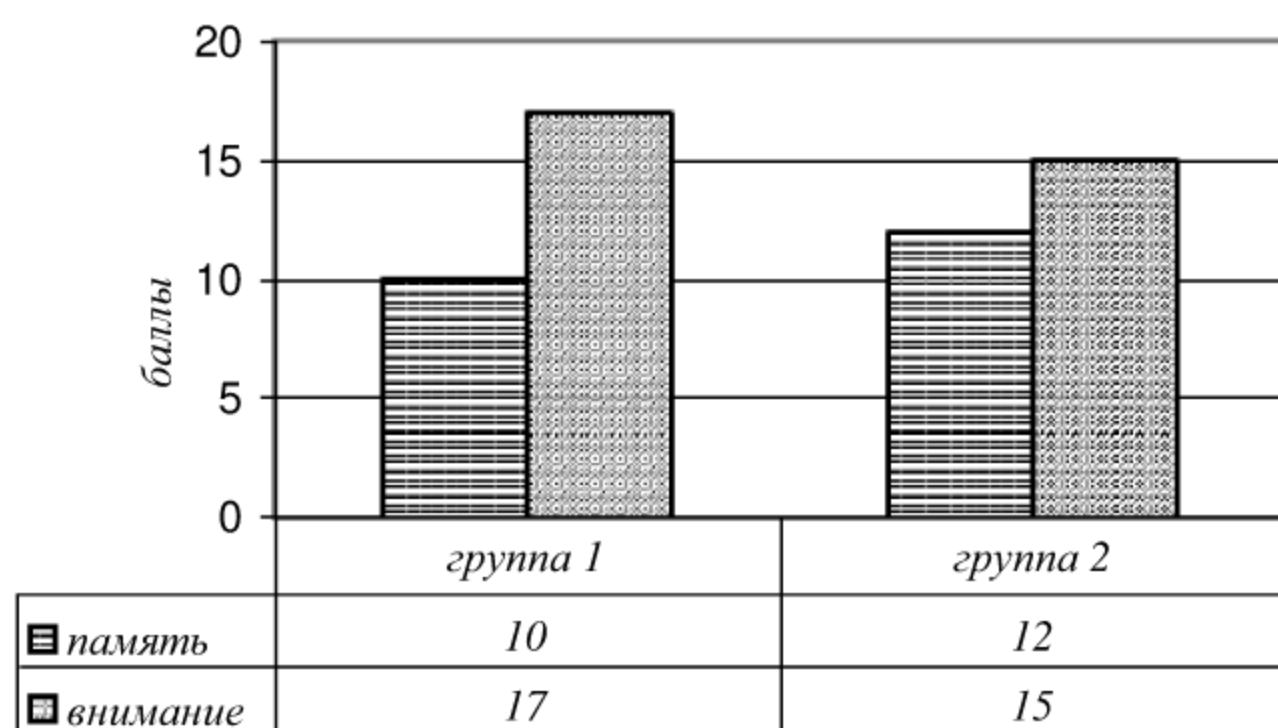
**Рис. 6.6. Правильно построенная столбиковая диаграмма различий**

При иллюстрации различий по нескольким показателям в группах следует располагать рядом столбики одного показателя (а еще лучше для иллюстрации различий по каждому показателю делать отдельный рисунок). Особенно это касается тех случаев, когда по-

казатели измеряются по разным методикам, у каждой из которых своя шкала измерения. Такого рода недочет (рис. 6.7), к сожалению, часто встречается при оформлении студенческих исследований (курсовых и дипломных работ).

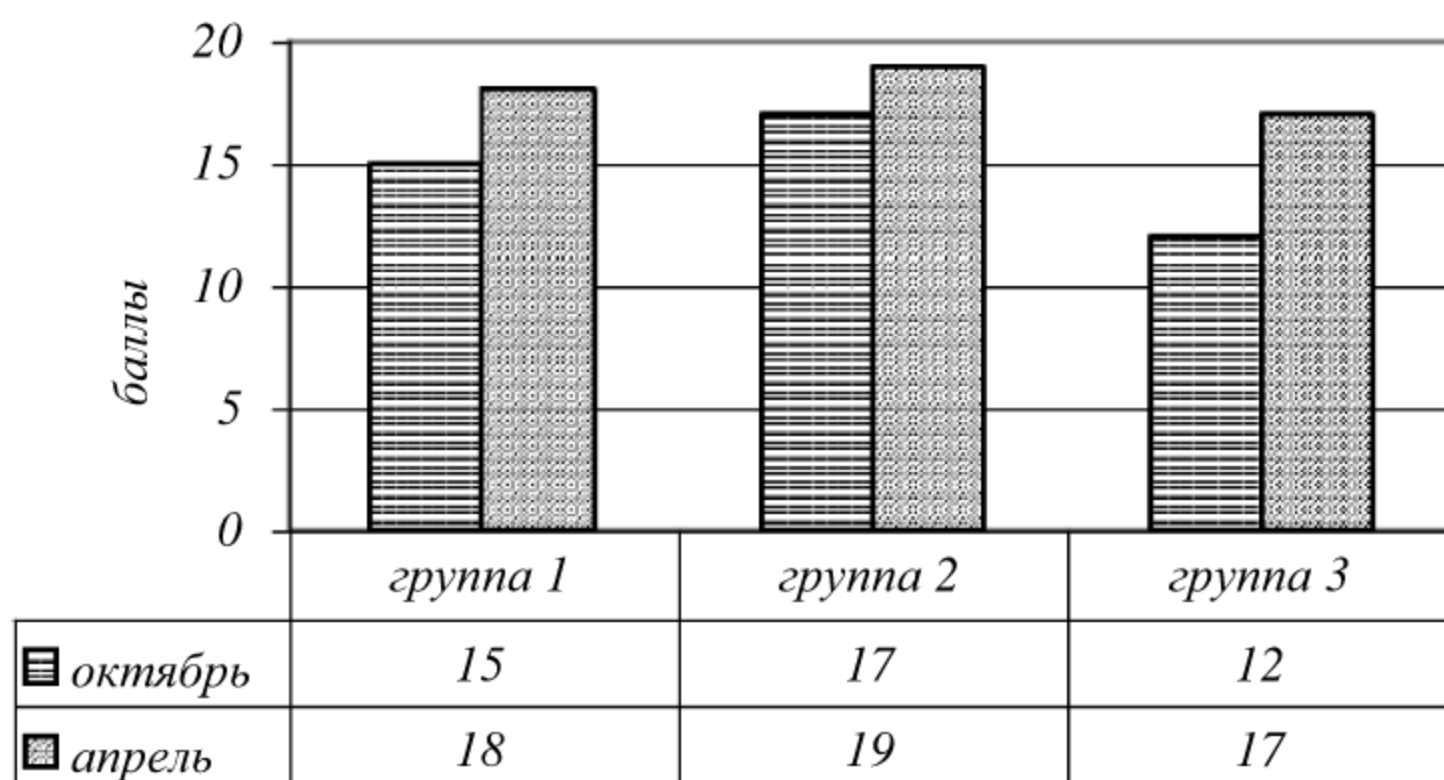
Столбиковые диаграммы используются и для иллюстрации изменений, т.е. различий двух рядов показателей, полученных на одних и тех же группах испытуемых с течением времени или при изменении каких-либо условий (например, в ситуациях покоя и стресса).

Техника построения столбиковой диаграммы изменений во многом совпадает с предыдущим случаем, но столбики диаграммы следует располагать по группам испытуемых.



**Рис.6.7. Неправильно построенная столбиковая диаграмма различий**

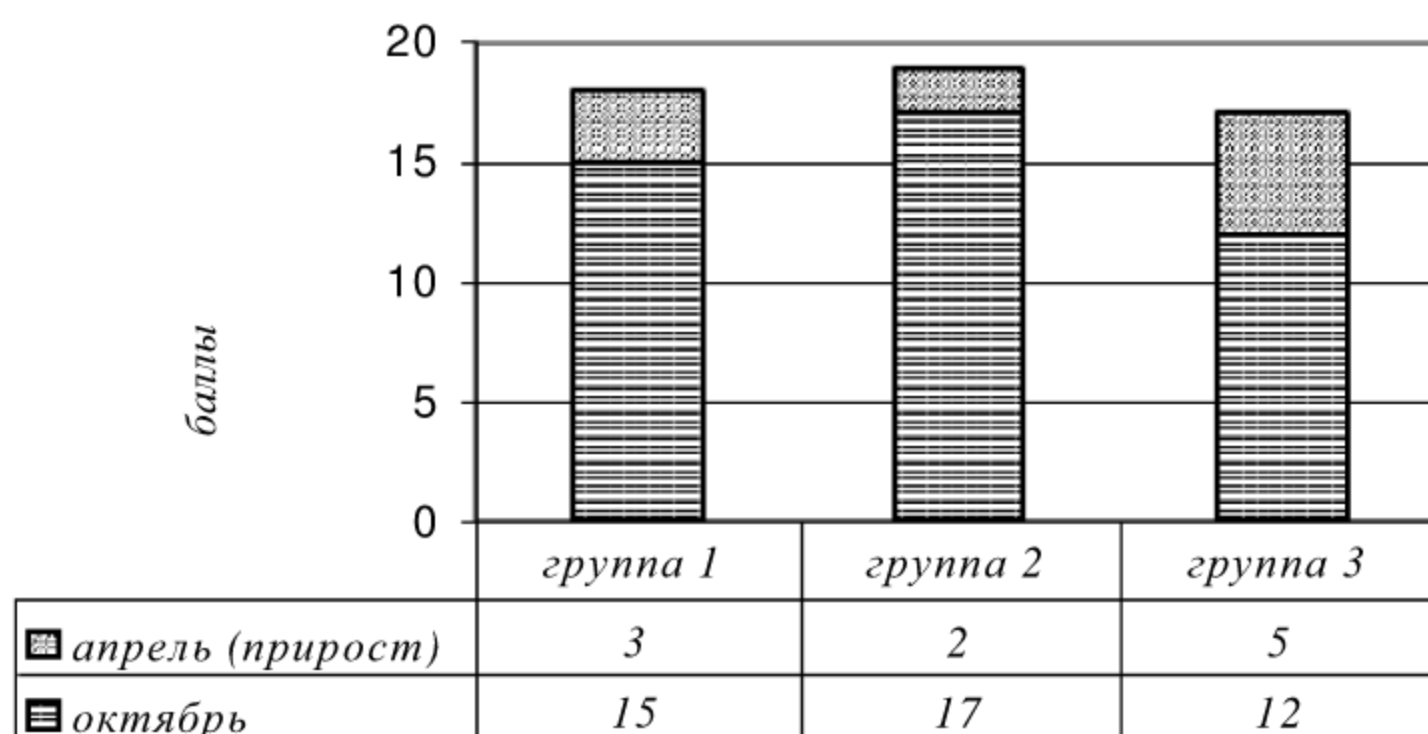
Рассмотрим пример. В трех группах учащихся вторых классов еженедельно по часу в течение пяти месяцев проводились занятия по развитию памяти. В ходе эксперимента изучалась эффективность занятий в каждой из групп. Для иллюстрации результатов можно использовать столбиковую диаграмму, изображенную на рис. 6.8. Высота столбиков соответствует уровню развития памяти в каждой из групп до и после занятий (в октябре и апреле).



**Рис.6.8 Столбиковая диаграмма изменений**

Сразу отметим, что столбиковая диаграмма изменений мало пригодна для иллюстрации результатов лонгитюдных исследований (три и более замеров). В этих случаях лучше использовать график, накладывая профили групповых изменений один на другой.

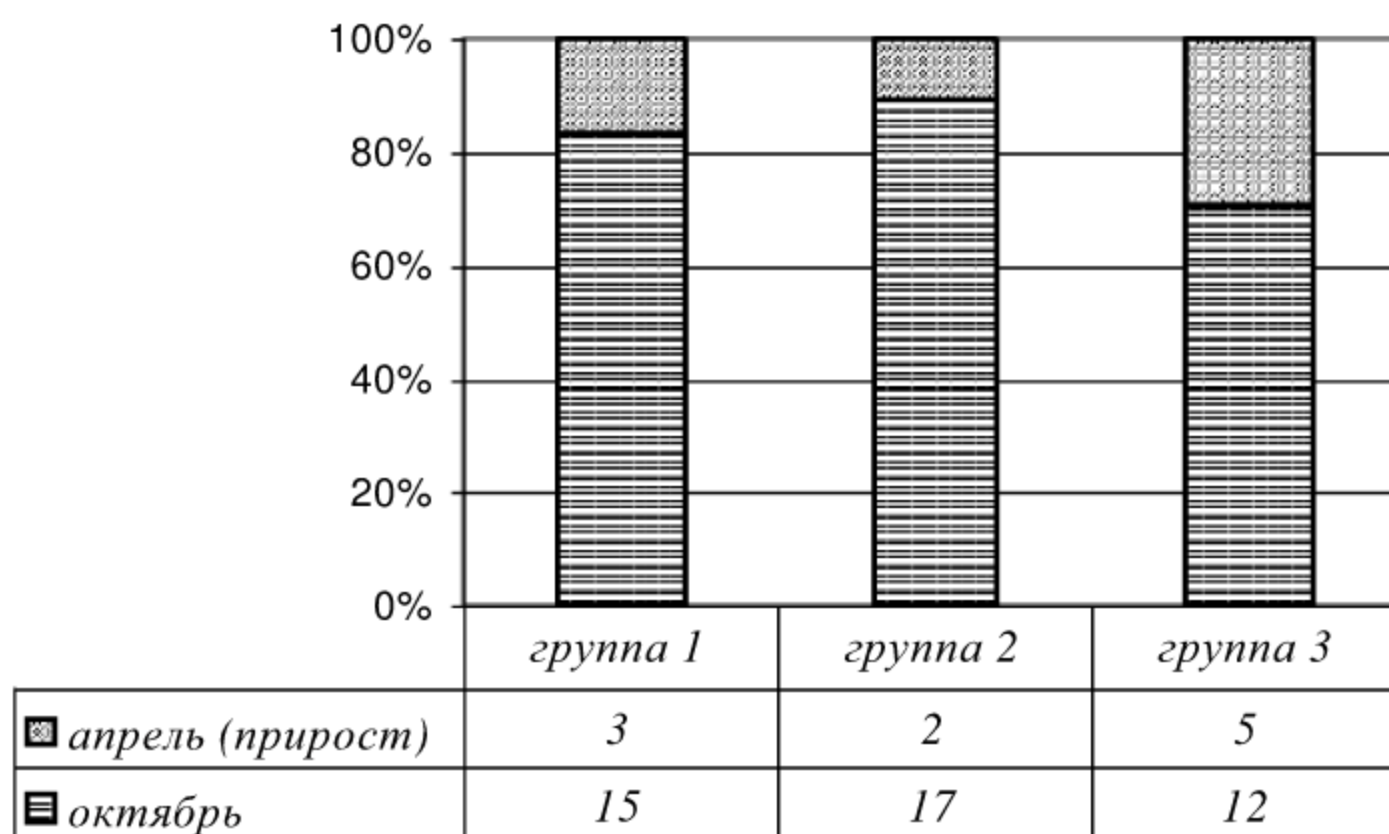
В ряде случаев удобно использовать столбиковые диаграммы с накоплением (обычные и нормированные), а также линейчатые диаграммы. Так, для наглядного изображения прироста показателя в каждой категории испытуемых с течением времени применяется столбиковая диаграмма с накоплением (рис. 6.9).



**Рис.6.9 Столбиковая диаграмма с накоплением**

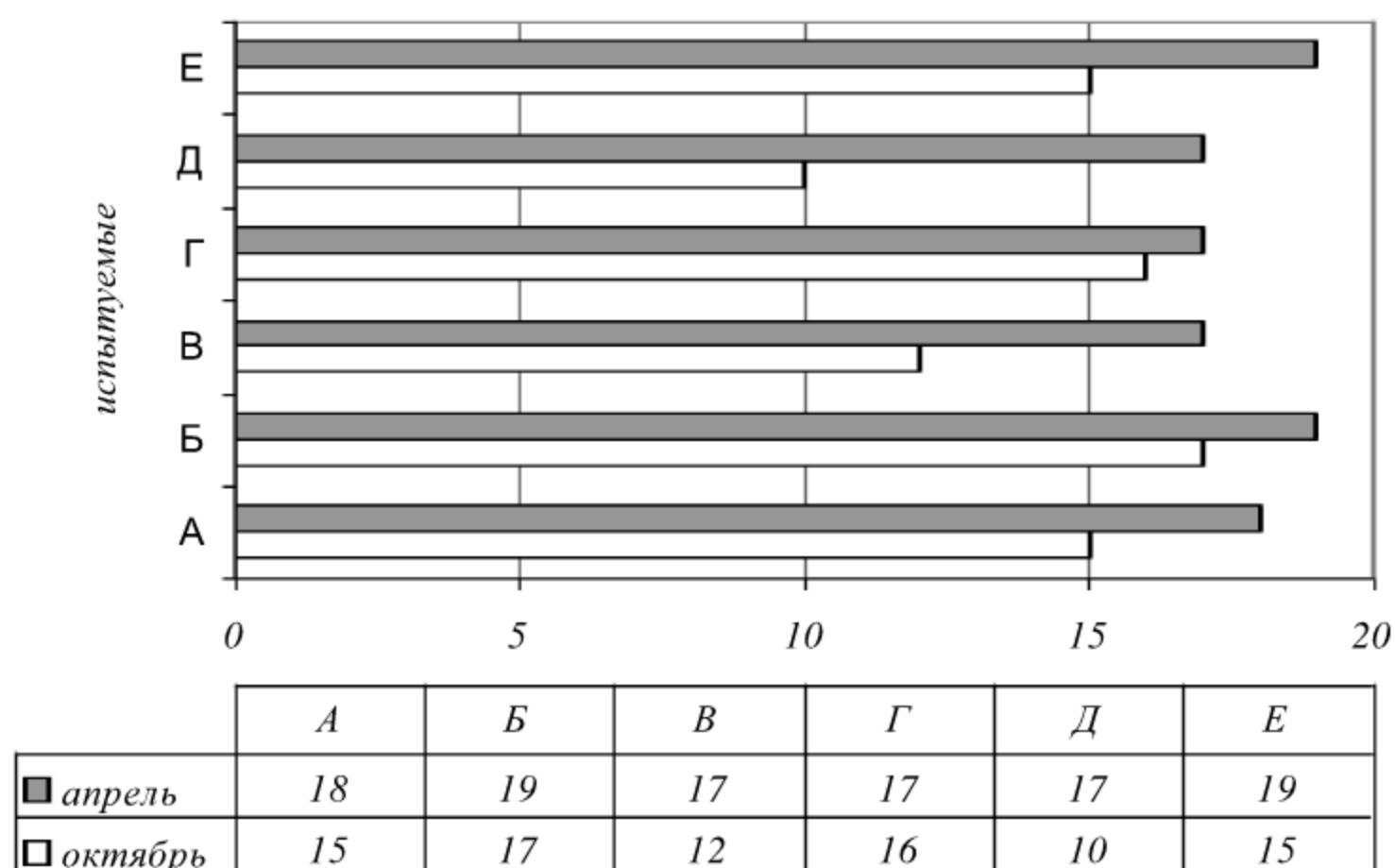


Если же требуется сделать акцент на сопоставление не абсолютных, а относительных величин изменений, то лучше использовать нормированную диаграмму с накоплением. На рис. 6.10 представлена иллюстрация к результатам эксперимента по развитию памяти в виде нормированной столбиковой диаграммы с накоплением, на которой видно, что наибольший относительный прирост уровня развития памяти наблюдается у школьников из третьей группы.

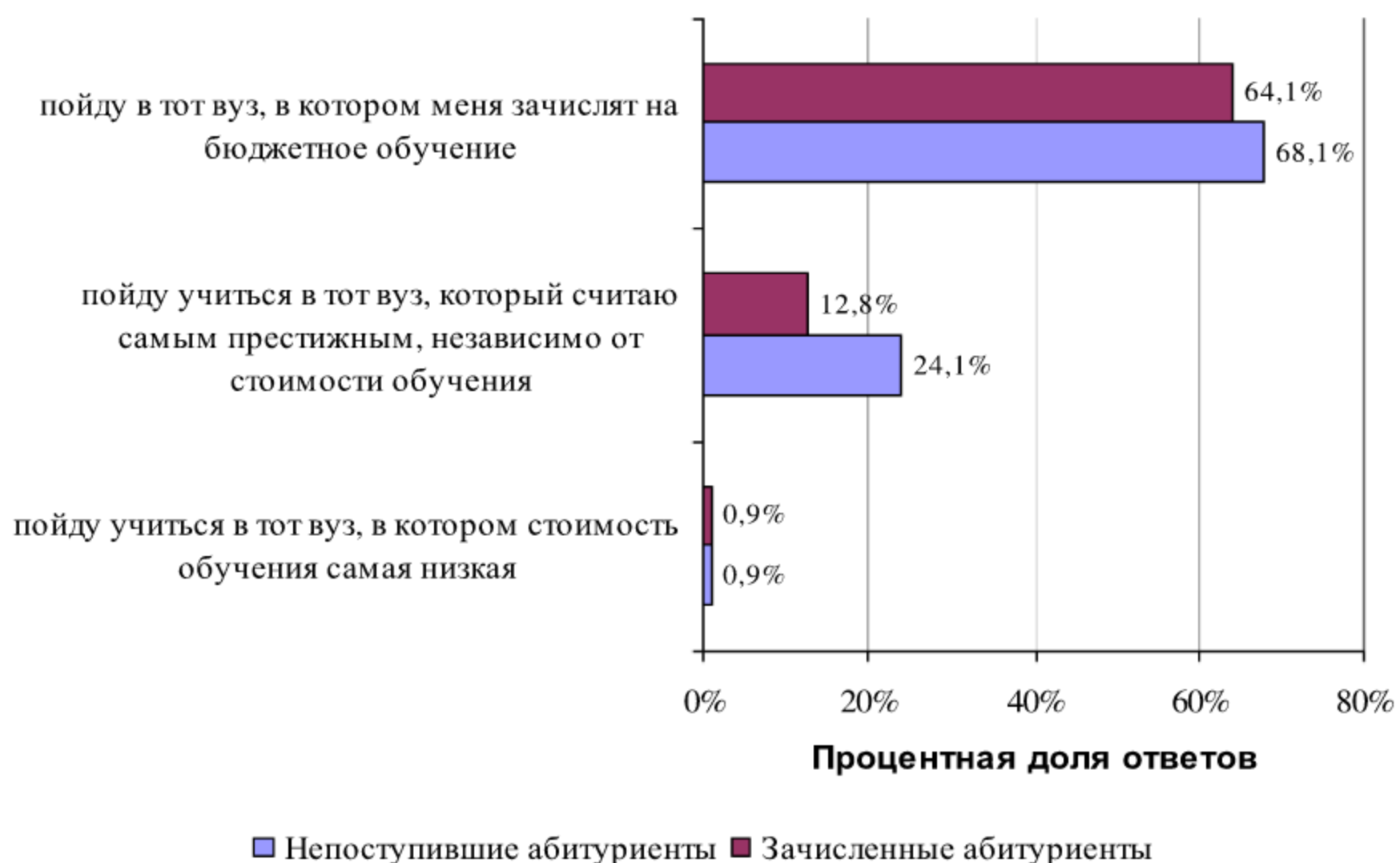


**Рис.6.10 Нормированная столбиковая диаграмма с накоплением**

В исследовательской работе психолога встречаются случаи, когда требуется представить иллюстрацию изменений индивидуальных результатов небольшой (до 10–15 человек) группы испытуемых. Для этой цели используются линейчатые диаграммы (рис. 6.11, а), принцип построения которых такой же, как и при построении столбиковых диаграмм. При достаточно большом количестве сопоставляемых данных узкие полоски линейчатой диаграммы визуально воспринимаются легче, чем слишком узкие вертикальные столбцы. Линейчатые диаграммы удобно использовать и для представления результатов анкетных опросов, образец такой диаграммы представлен на рис. 6.11 б).



а) линейчатая диаграмма изменений показателей в небольшой выборке испытуемых (6 человек)



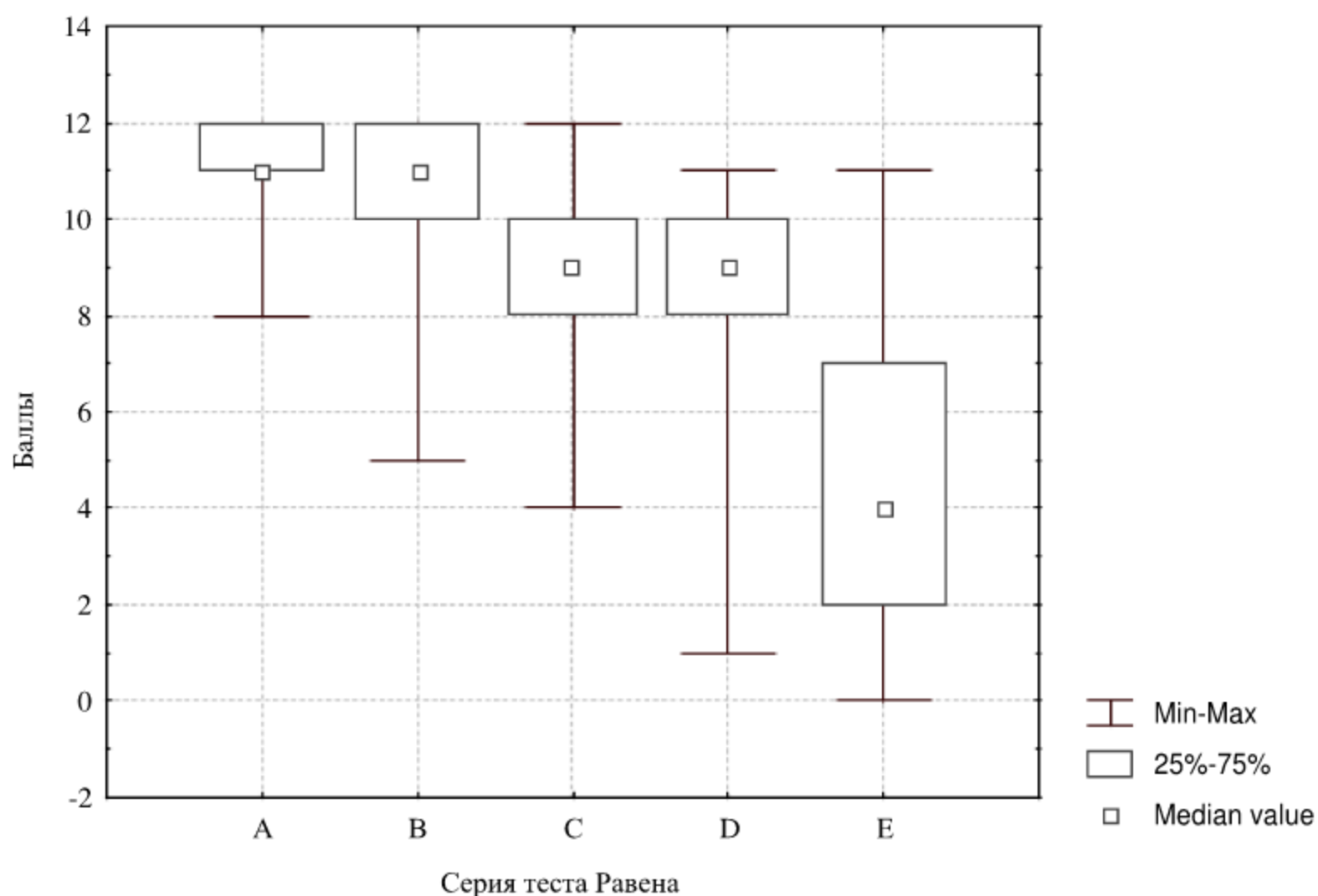
б) Распределение ответов зачисленных и непоступивших абитуриентов на вопрос «Если Вас примут сразу в несколько вузов, что определит Ваш окончательный выбор?»

Рис.6.11 Образцы линейчатых диаграмм



## Диаграмма размаха

Диаграммы размаха (называемые графиками «ящички-усы») применяются для визуального сопоставления диапазонов и вариаций значений переменной в различных группах испытуемых. Для их построения в каждой группе вычисляются квартили. Центр «ящичка» соответствует медиане (или  $Q_2$ ), границы «ящичка» определяются значениями  $Q_1$  и  $Q_3$ , а границы «усов» – максимальным и минимальным значениями показателя в соответствующей группе испытуемых. Диаграмму размаха часто используют и для иллюстрации результатов обследования одной группы испытуемых с помощью многомерной методики (например, тестов Векслера, Равена, 16ФЛО), когда результаты по каждому показателю или субтесту измеряются по одинаковым шкалам. На рис. 6.12 изображена диаграмма размаха, иллюстрирующая диапазоны результатов выполнения серий теста Равена учащимися третьих классов.



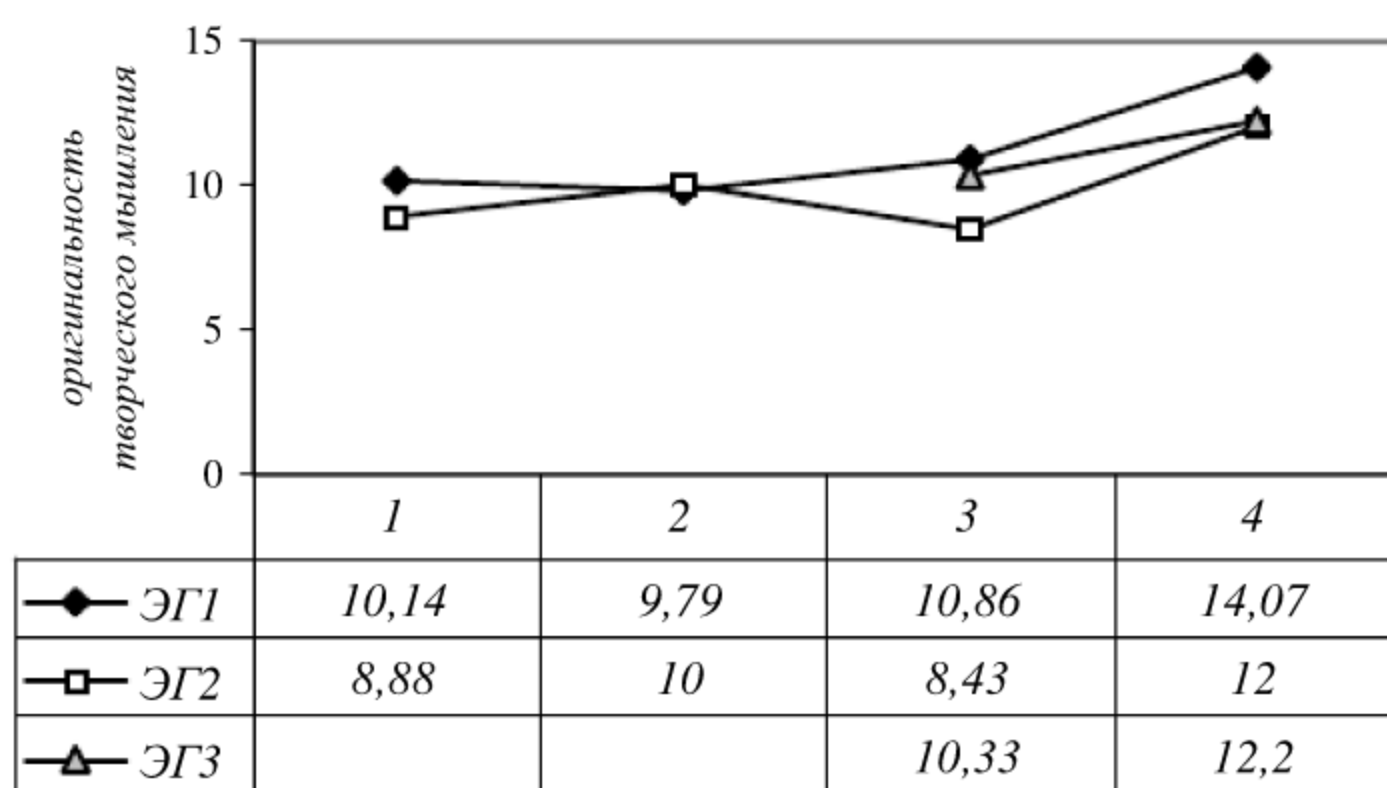
**Рис. 6.12. Диаграмма размаха**

Отметим, что при построении диаграммы размаха выбор меры положения (центр «ящичка») и меры вариативности признака (гра-

ницы «ящика») производится в соответствии со шкалой измерения признака. Так, для порядковой шкалы целесообразно использование квартилей, а для интервальной центр «ящика» может определяться средним значением, а его границы – стандартным отклонением, вычисленными для каждой группы наблюдений.

### График

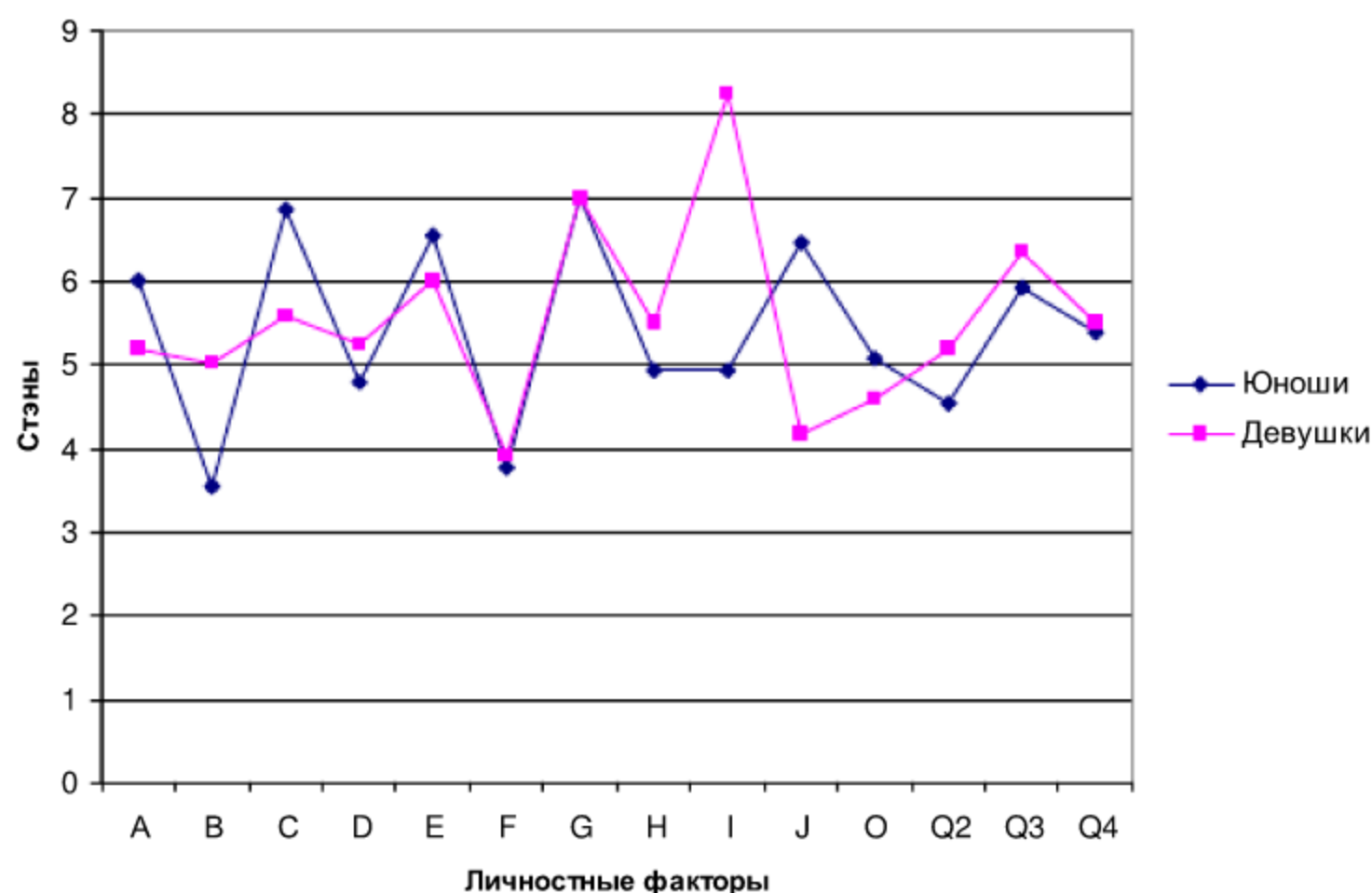
Графики используются чаще всего для изображения сравнительной динамики изменения одного показателя у различных типов или групп испытуемых (на *рис. 6.13* представлен график сравнительной динамики оригинальности творческого мышления трех экспериментальных групп школьников, принимавших участие в лонгитюдном исследовании, в ходе которого было произведено четыре замера). Во многих случаях график иллюстрирует выявленные тенденции в изменении изучаемого признака.



**Рис. 6.13 График сравнительной динамики развития творческого мышления в экспериментальных группах школьников**

Графики также традиционно используются для представления личностных профилей у испытуемых, а также групповых профилей у различных групп испытуемых (рис. 6.14). При анализе групповых профилей следует иметь в виду, что значения показателей группового профиля, близкие к средним, могут свидетельствовать о разбросе данного показателя (у одной части группы, например, встре-

чаются высокие значения показателя, а у другой части – низкие, которые на групповом профиле «компенсируются»).



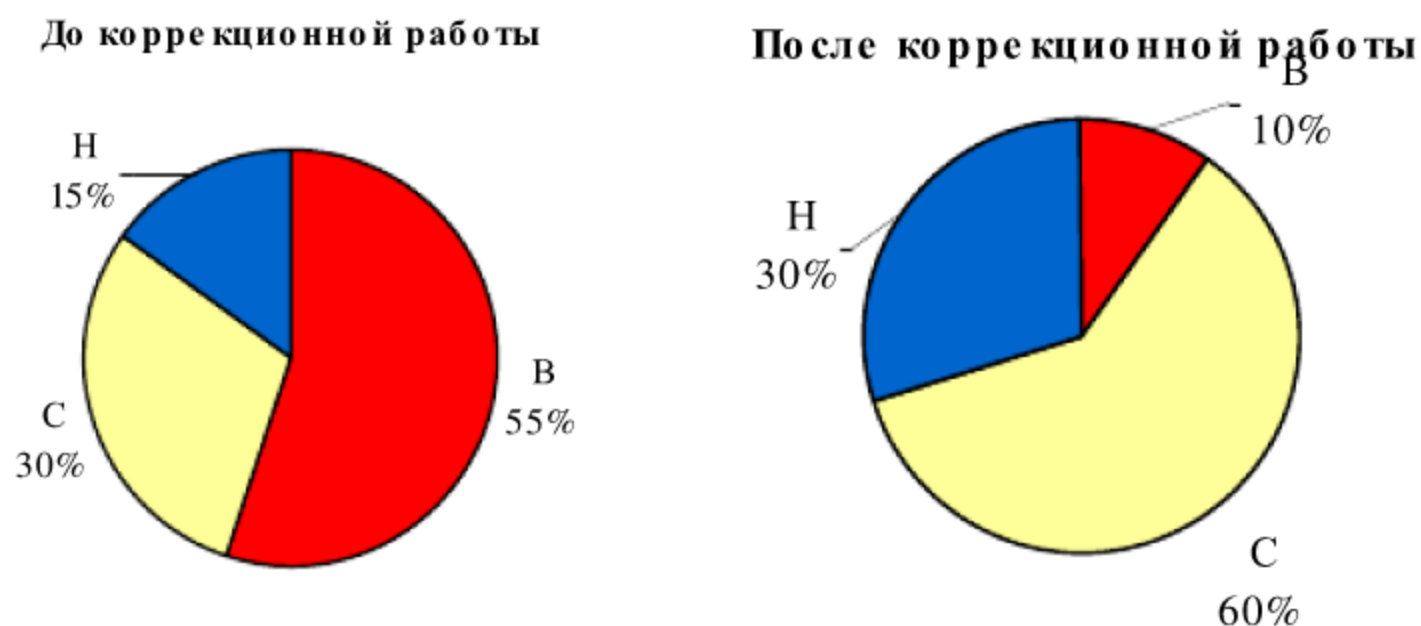
**Рис. 6.14 Личностные профили юношей и девушек**

Если же на групповом профиле наблюдаются высокие или низкие значения показателя, то этот факт, скорее всего, свидетельствует о соответствующей тенденции, статистическую значимость которой необходимо проверить с помощью статистических критериев.

### Круговая диаграмма

Круговые диаграммы применяются для наглядного представления процентного соотношения уровней измеренного признака в группе испытуемых. Для построения круговой диаграммы нужно определить число испытуемых, получивших низкую, среднюю и высокую оценку измеряемого свойства, затем вычислить процентную долю каждой группы в общей обследованной выборке и соответствующую величину угла сектора диаграммы.

На рисунке 6.15 представлены круговые диаграммы, иллюстрирующие процентное соотношение испытуемых с высоким, средним и низким уровнем тревожности до и после тренинга эмоциональной саморегуляции.

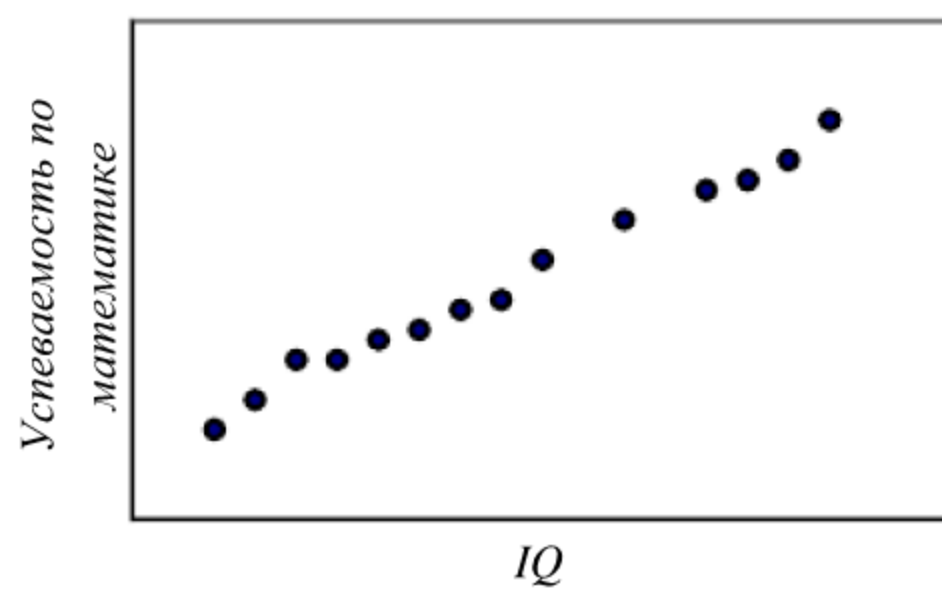


**Рис. 6.15 Уровень тревожности в группе до и после тренинга эмоциональной саморегуляции**

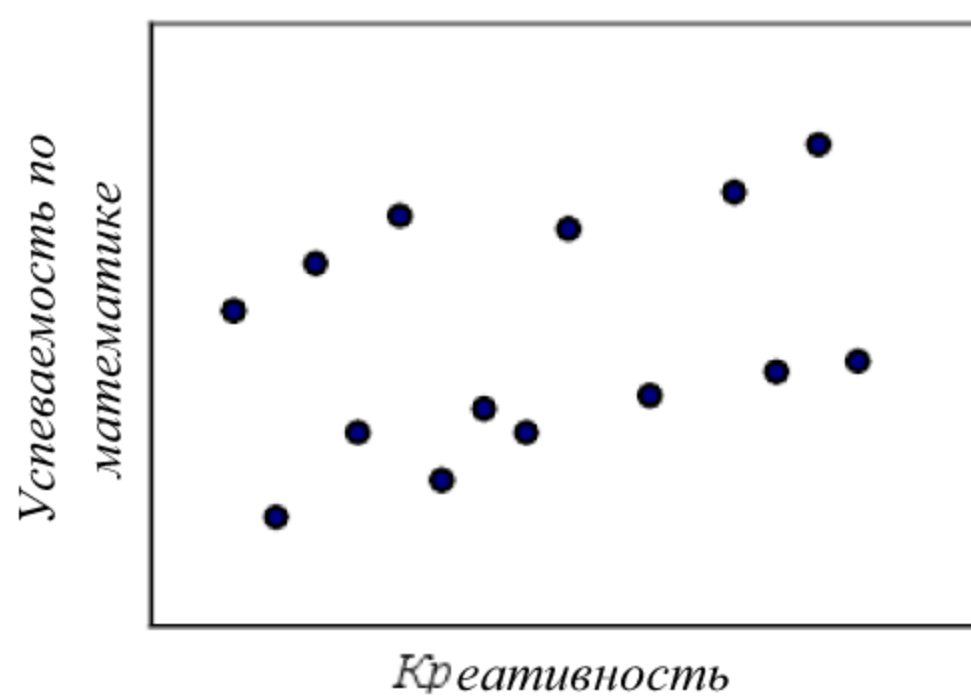
### Точечная диаграмма

Точечная диаграмма (диаграмма рассеяния) используется для иллюстрации наличия или отсутствия связи между двумя измеренными признаками в выборке испытуемых (например, IQ и креативность у студентов). Для ее построения по осям откладываются значения признаков, охватывающие весь диапазон значений в выборке. Оценки каждого испытуемого в выборке отмечаются точкой, абсцисса и ордината которой соответствуют значениям признаков у этого испытуемого. В результате число точек на диаграмме равно объему выборки, а их взаимное расположение характеризует степень согласованности изменений измеренных признаков.

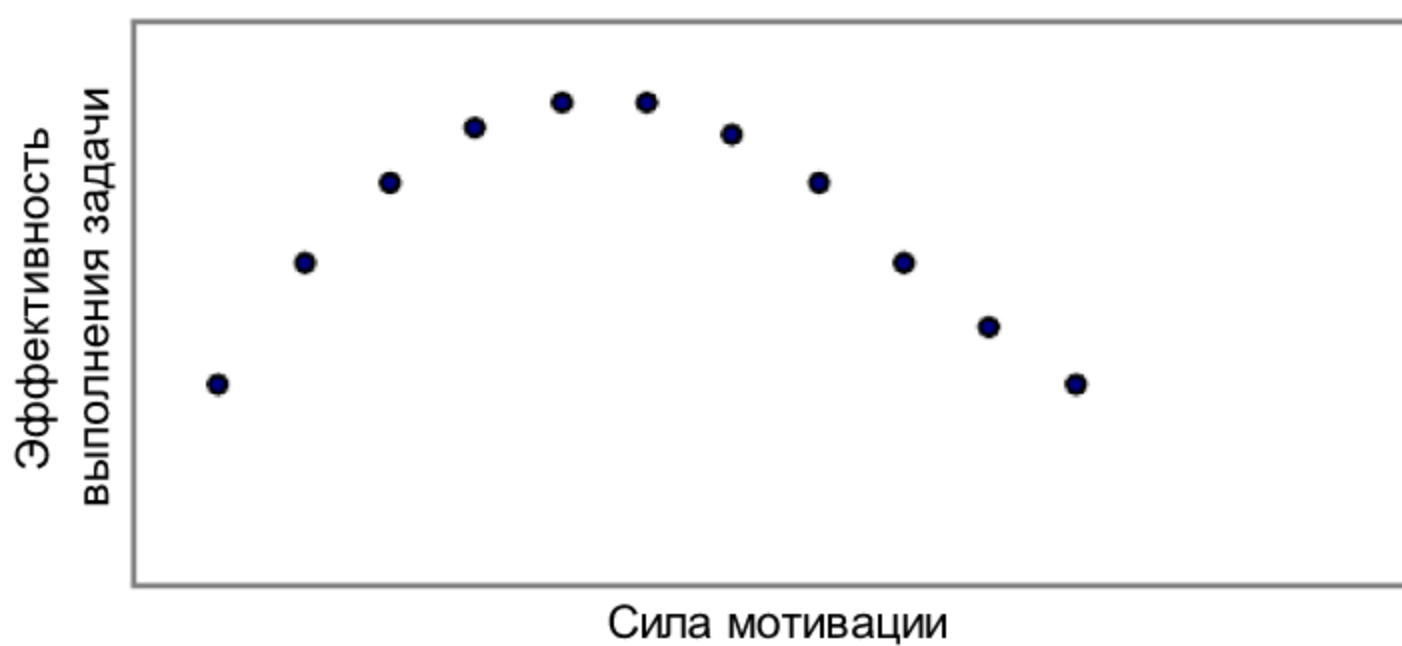
На рис.6.16 изображена диаграмма рассеяния, иллюстрирующая положительную линейную связь между параметрами (IQ – коэффициентом интеллекта и успеваемостью по математике у младших школьников). На рис.6.17 точки расположены хаотично, что иллюстрирует отсутствие связи между переменными (креативностью и успеваемостью по математике). Отметим, что в некоторых случаях построение диаграммы рассеяния позволяет обнаружить нелинейную связь между переменными: такова, например, связь между силой мотивации и эффективностью выполнения задачи (рис.6.18).



**Рис.6.16 Положительная линейная связь между переменными**



**Рис. 6.17 Отсутствие связи между переменными**

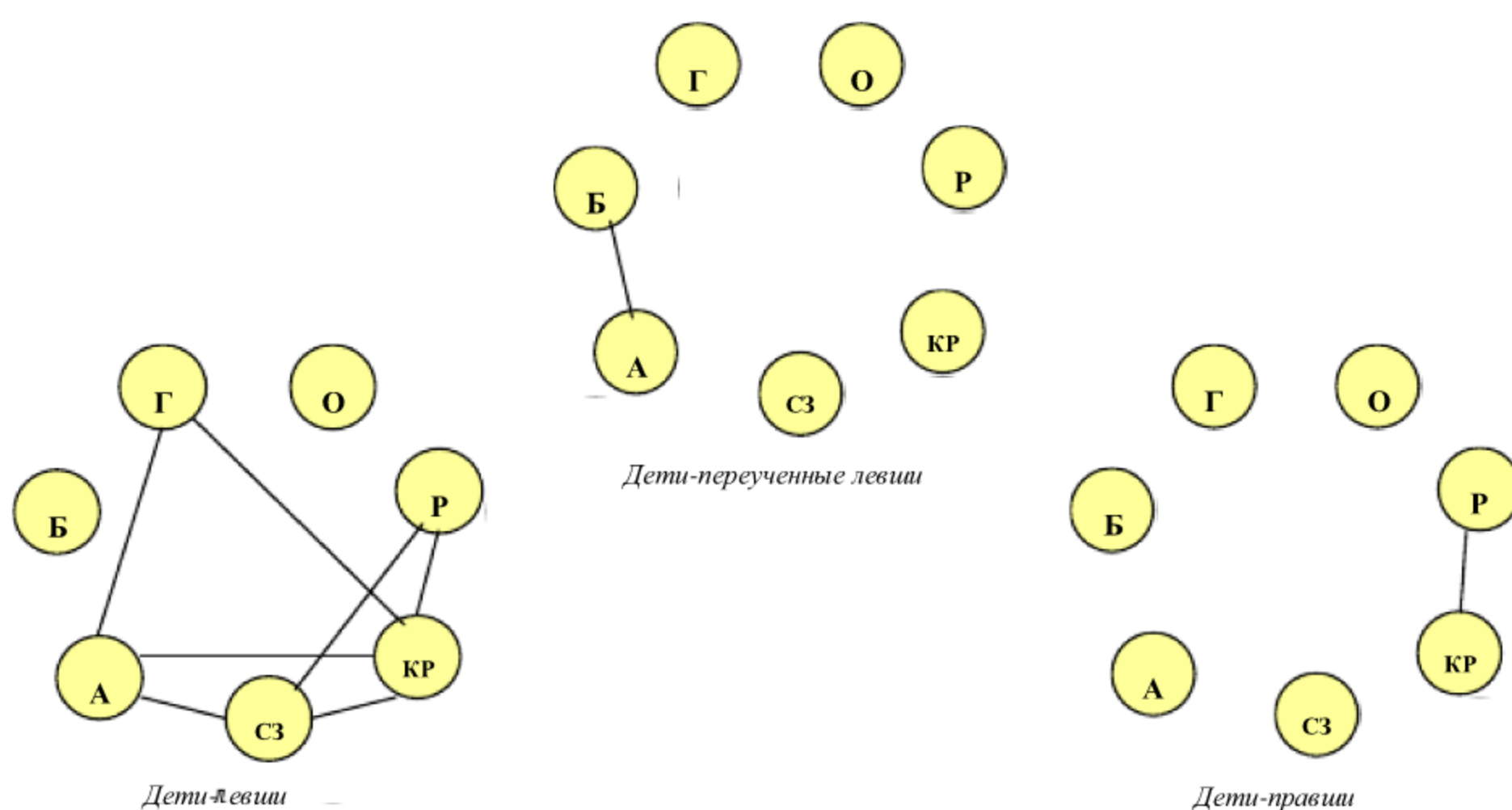


**Рис. 6.18. Связь между силой мотивации и эффективностью выполнения задачи**



## Корреляционный граф

Для иллюстрации взаимосвязей внутри множества показателей, полученных при обследовании выборки испытуемых используются корреляционные графы. Корреляционные графы, построенные для нескольких сопоставляемых выборок, характеризуют структурные различия между группами по множеству измеренных показателей. На рисунке 6.19 представлены корреляционные графы, иллюстрирующие взаимосвязь между показателями творческого мышления испытуемых с различным мануальным предпочтением.



**Рис. 6.19. Связь показателей креативности в группах испытуемых в зависимости от мануального предпочтения**

## Контрольные вопросы и задания

Определите шкалу измерения:

1. время (в секундах) выполнения теста Д.Равена;
2. образование испытуемых (среднее, бакалавр, магистр);
3. IQ в тесте Д.Векслера;
4. тип темперамента в опроснике Г.Айзенка (EPQ);
6. число верно выполненных заданий в тесте Д.Равена;



6. *число верно выполненных заданий в тесте Д.Равена, как показатель интеллектуального развития;*
7. *пол испытуемых;*
8. *возраст испытуемых;*
9. *ФИО испытуемых;*
10. *умственный возраст в тесте А.Бине.*

### **6.3. Числовые характеристики множества данных**

Множество данных, полученных в результате психодиагностического обследования выборки, можно охарактеризовать с помощью двух показателей, первый из которых является мерой центральной тенденции и определяет типичное значение для данной выборки, а второй является мерой изменчивости и определяет разброс значений. Во многих случаях такая характеристика результатов вполне приемлема, хотя при ее использовании часть информации утрачивается.

#### **Меры центральной тенденции (меры положения)**

Меры центральной тенденции служат для определения «типичных значений» или «центрального положения» совокупности данных. При выборе меры центральной тенденции, способной наиболее точно охарактеризовать выборку, следует учитывать, во-первых, тип шкалы, в которой производились измерения, и, во-вторых, особенности самой выборки.

**Мода (Мо)** – это наиболее часто встречающееся значение в множестве наблюдений. Иными словами, это значение, имеющее наибольшую вероятность появления. Так, в совокупности значений (3, 5, 5, 7, 9, 9, 9, 10) модой является число 9. Для данных, измеренных по шкале наименований, мода – единственная мера центральной тенденции. Так, например, если в группе 10 сангвиников, 7 холериков, 5 флегматиков и 2 меланхолика, то модой (наиболее часто встречающимся темпераментом в группе) будет сангвиник. Не вся-

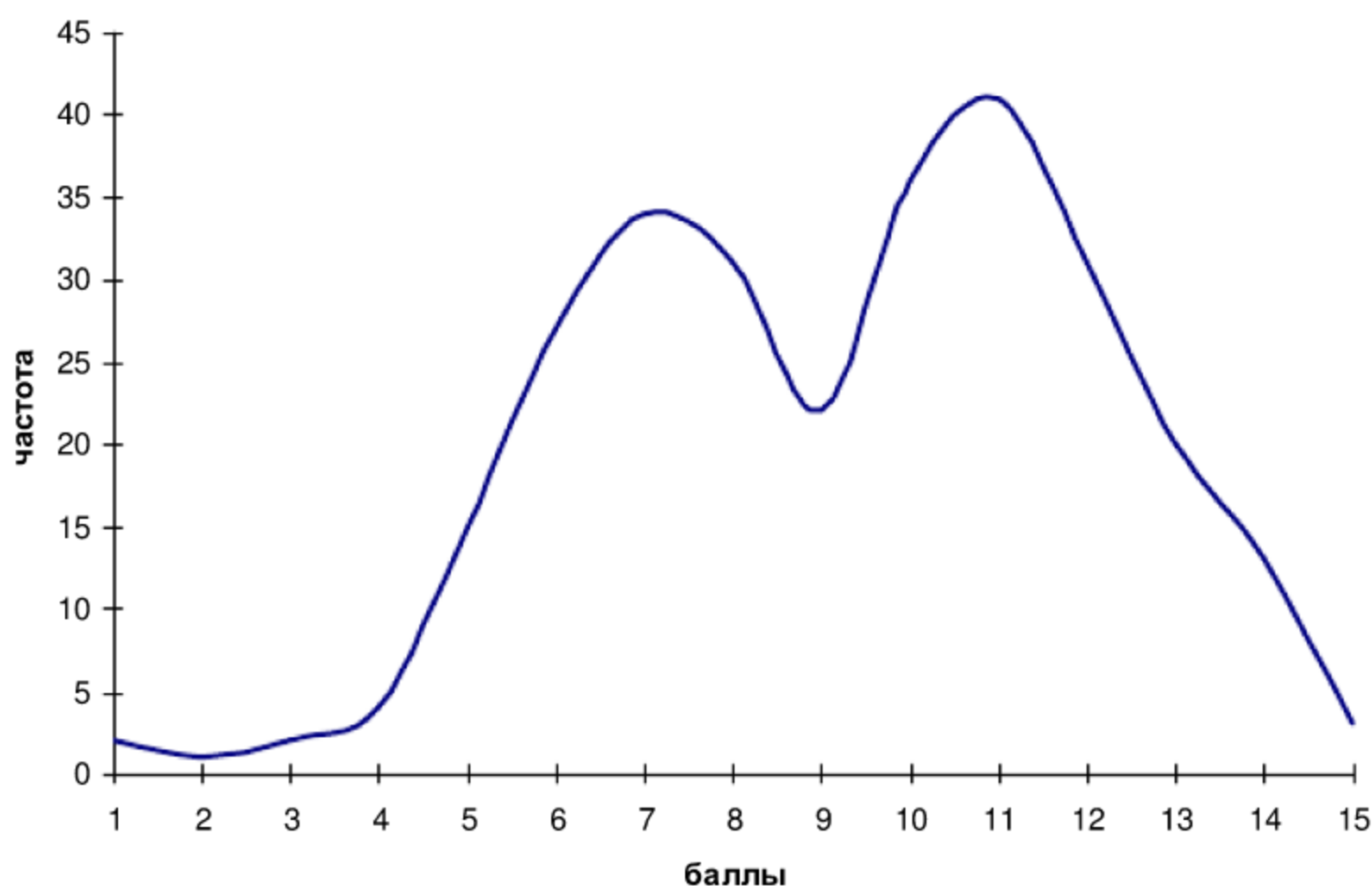
кая совокупность значений имеет единственную моду, поэтому существуют следующие правила определения моды.

В случае, когда все значения в группе встречаются одинаково часто, принято считать, что группа оценок не имеет моды. Так, в группе значений (2, 2, 3, 3, 4, 4) моды нет.

Если данные измерены в порядковой или интервальной шкале, и два соседних числовых значения имеют одинаковую частоту, которая больше частоты любого другого значения, то мода вычисляется как среднее арифметическое этих двух значений. Так, мода группы значений (0, 1, 2, 2, 3, 3, 3, 4, 4, 4, 5, 5) равна  $Mo = (3+4)/2 = 3,6$ .

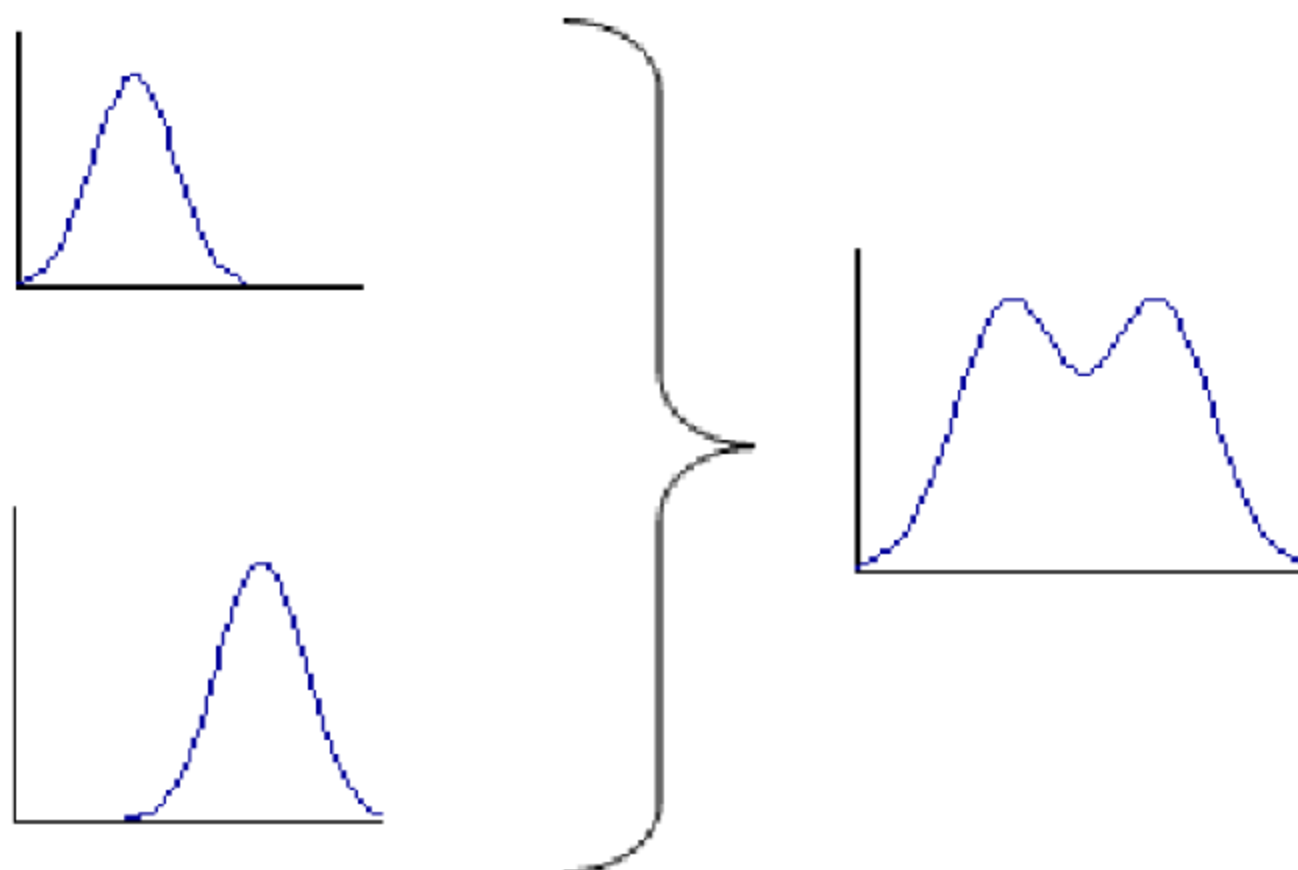
Если два несмежных значения в группе имеют равные частоты и они больше частот любого значения, то существуют две моды; в таком случае группа оценок называется бимодальной. Например, в группе значений (10, 11, 11, 11, 12, 13, 14, 14, 14, 16) модами являются 11 и 14.

Большие множества данных называются бимодальными, когда они образуют полигон частот, похожий на спину двугорбого верблюда, даже если частоты на двух вершинах не строго равны (рис. 6.20).



**Рис. 6.20. Бимодальное распределение**

Аналогичным образом определяются мультимодальные распределения. Как правило, бимодальные и мультимодальные множества данных возникают вследствие неоднородности обследованной выборки и являются объединением соответствующего числа унимодальных множеств (рис. 6.21).



**Рис. 6.21. Бимодальное распределение, как объединение унимодальных распределений**

Как уже отмечалось выше, в качестве меры центральной тенденции множества значений признака, измеренного по номинативной шкале, можно использовать только моду. Моду также удобно использовать при группировке данных, измеренных по порядковой и интервальной шкале. Например, по результатам обследования группы испытуемых можно подсчитать число респондентов, получивших в соответствии со статистическими нормами высокое, среднее и низкое значения измеряемого показателя, и соответственно его модальное значение. В этих случаях полученный результат может быть проиллюстрирован с помощью круговой диаграммы. Отметим, что в приведенном примере шкала измерения признака была понижена, и, следовательно, утрачена часть информа-

ции. Вместе с тем, подобная операция зачастую позволяет повысить «читаемость» полученных результатов.

**Медиана (Me)** – это значение, которое делит упорядоченное множество значений параметра пополам, так что одна половина значений оказывается больше медианы, а другая – меньше. Другими словами, медиана представляет собой 50-й процентиль в группе данных

$$Me = P_{50} \quad (6.2)$$

Для вычисления медианы множество значений признака, полученных при обследовании выборки испытуемых, нужно упорядочить.

Пусть  $X_1, X_2, \dots, X_n$  – упорядоченное множество значений признака. Тогда возможны два случая:

1).  $n$  – нечетное, т.е.  $n = 2k + 1$ .

В этом случае медиана вычисляется по формуле:

$$Me = X_{k+1}; \quad (6.3)$$

2).  $n$  – четное, т.е.  $n = 2k$ .

В этом случае медиана вычисляется по формуле:

$$Me = \frac{X_k + X_{k+1}}{2}. \quad (6.4)$$

Например, в множестве значений 10, 11, 11, 11, 12, 13, 14, 14, 16 ( $n = 9 = 2 \cdot 4 + 1$  – нечетное,  $k = 4$ ) медиана  $Me = X_5 = 12$ , а в множестве значений 9, 10, 11, 11, 11, 14, 14, 15, 15, 16 ( $n = 10 = 2 \cdot 5$  – четное,  $k = 5$ ) медиана

$$Me = \frac{X_5 + X_6}{2} = \frac{11 + 14}{2} = 12,5. \quad (6.5)$$

Как правило, медиана используется в качестве меры центральной тенденции множества данных, измеренных по порядковой шкале, и особенно удобна в тех случаях, когда отдельные значения показателя существенно превышают подавляющее большинство значений показателя в выборке. Характерным свойством медианы является ее нечувствительность к «скачкам» на концах измери-

тельной шкалы. Так, значение медианы не изменится в случае, когда наибольшее значение признака, например, утроится. В множестве значений 1, 11, 11, 11, 12, 13, 14, 14, 48 ( $n = 9 = 2 \cdot 4 + 1$  – нечетное,  $k = 4$ ) медиана  $Me = X_5 = 12$ , как и в примере, рассмотренном выше.

Таким образом, на медиану не влияют величины «больших» и «малых» значений, следовательно, она характеризует наиболее типичные значения признака в выборке испытуемых. Однако медиану в качестве меры центральной тенденции множества данных, содержащих крайние значения, следует использовать только в тех случаях, когда распределение унимодально либо не имеет моды. Мультимодальные распределения чаще всего являются следствием неоднородности выборки и, как правило, не подлежат описанию с помощью одного-двух числовых значений, т.к. как не имеют центральной тенденции. Пример такого распределения будет рассмотрен ниже.

**Среднее арифметическое (M)** совокупности  $n$  значений  $X_1, X_2, \dots, X_n$  определяется по формуле 6.6.

$$M = \frac{X_1 + X_2 + \dots + X_n}{n} = \frac{1}{n} \sum_{i=1}^n X_i. \quad (6.6)$$

Кроме среднего арифметического существуют и другие средние значения: среднее гармоническое, среднее геометрическое, среднее квадратическое, среднее кубическое и т.д., но при представлении результатов психодихологических исследований они используются редко. Основной мерой положения является среднее арифметическое значение. Средние значения второй, третьей и четвертой степеней используются при вычислении мер рассеяния, асимметрии и эксцесса.

Как правило, среднее арифметическое значение как меру центральной тенденции следует использовать в случаях распределения данных, близкого к нормальному, т.е. когда большинство значений признака сгруппировано вокруг некоторого центрального значения,

положительные и отрицательные отклонения от которого практически равновероятны и частота их уменьшается с удалением от центра. Заметим, что в случае такого распределения данных значения моды, медианы и среднего отличаются незначительно, а если распределение унимодально и строго симметрично, то совпадают. В других случаях, особенно когда имеются выбросы на концах шкалы, к которым среднее арифметическое значение наиболее чувствительно, лучше использовать медиану.

На величину среднего влияет каждое значение. Одно крайнее значение может очень сильно сместить среднее группы. Например, если 9 человек имеют доходы от 2500 до 5000 рублей в месяц со средним 4000 рублей, а доход десятого составляет 30000 рублей, то средний ежемесячный доход для 10 лиц будет составлять 6600 рублей. Эта цифра совершенно не отражает реального положения дел.

Рассмотрим пример множества данных, не имеющего центральной тенденции. Группа студентов выполнила контрольную работу, состоящую из 8 заданий по изученной теме. За каждое правильно выполненное задание начислялся 1 балл. Результаты обследования представлены в табл. 6.13 и на гистограмме распределения частот (рис. 6.22).

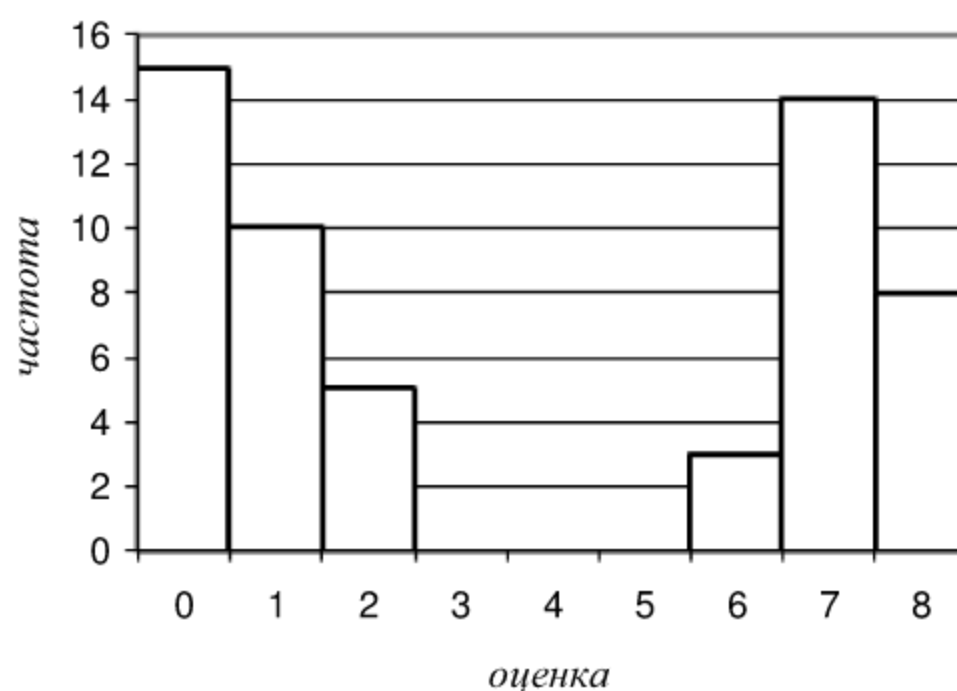
**Таблица 6.13. Распределение оценок в группе**

Оценка	Частота
0	15
1	10
2	5
6	3
7	14
8	8

Средняя оценка в группе  $M=3,64$ , а медиана  $Me=4$ . Ни одно из этих значений не несет никакой информации о полученных результатах, в первую очередь, о том, что распределение оценок в группе бимодально ( $Mo_1=0$ ,  $Mo_2=7$ ), т.е. группа фактически разделилась на «успевающих», получивших 6, 7 и 8 баллов, и «неуспевающих», получивших 0, 1 и 2 балла. Заметим, что значения среднего и ме-



дианы были бы точно такими же, если бы результаты контрольного испытания оказались бы более ровными и 18 студентов получили оценку 3, а 32 – оценку 4.



**Рис.6.22. Гистограмма распределения оценок**

Таким образом, прежде чем вычислять среднее значение или медиану множества значений, необходимо убедиться в существовании центральной тенденции данного множества. Для этого в большинстве случаев достаточно построить гистограмму распределения частот. В студенческих курсовых и дипломных работах, к сожалению, приходится встречать отступления от этого правила, что приводит авторов исследований к неверным выводам.

Рассмотрим еще один пример (табл. 6.14).

**Таблица 6.14 Результаты обследования двух групп испытуемых.**

№	1 группа	2 группа
1	19	23
2	21	24
3	23	25
4	24	26
5	25	27
6	25	27
7	27	31
8	53	34
Me	24,5	26,5
M	27,125	27,125

Психолог обследовал интеллектуальное развитие двух групп первоклассников с помощью стандартных прогрессивных матриц Д.Равена. В каждой группе было 8 испытуемых. В таблице 6.14 представлены результаты обследования, а также среднее значение и медиана по каждой группе. Как показывает визуальный анализ таблицы, средние значения в обеих выборках совпадают. Вместе с тем, во второй группе у большинства испытуемых оценки выше (и, как следствие, выше значение медианы), а в первой группе больше разброс значений. Таким образом, для более полной характеристики выборки помимо меры центральной тенденции необходимо знание меры изменчивости (вариативности) данных.

### Меры изменчивости

Если меры центральной тенденции говорят о концентрации группы значений на числовой шкале, то меры изменчивости дают общее представление о вариации оценок внутри группы.

**Размах** – это разность максимального и минимального значений в группе:

$$d = x_{\max} - x_{\min} \quad (6.7)$$

Размах вычисляется очень просто и является общераспространенной мерой изменчивости. Вместе с тем, он очень чувствителен к отдельным выбросам на концах шкалы. Уникальные результаты одного испытуемого могут значительно увеличить размах. Как видно из *табл. 6.13*, размах значений в первой группе значительно (в 3 раза) выше, чем во второй, только за счет единственного испытуемого, набравшего 53 балла. Немного отойдя от описательной статистики, отметим, что такой результат сам по себе уникален для первоклассника и должен направить психолога на дополнительные обследования ребенка, а также на работу с его учителями и родителями.

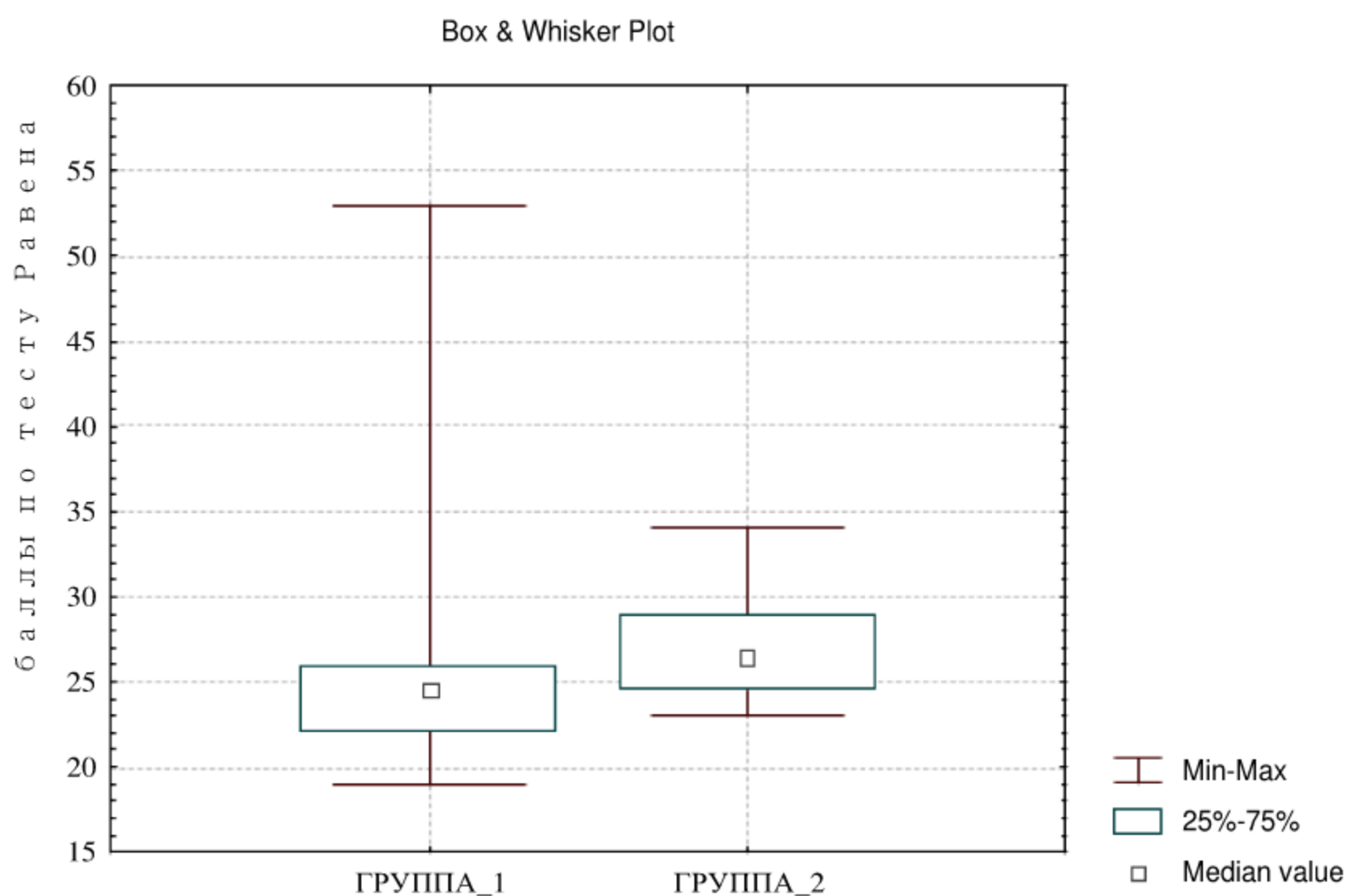
Более стабильными характеристиками вариативности значений в группе являются размах между 10-м и 90-м процентилями в группе значений

$$D = P_{90} - P_{10}, \quad (6.8)$$

а также квартильный размах

$$Q = Q_3 - Q_1. \quad (6.9)$$

Легко заметить, что квартильный размах – это интервал, в который попадает 50% наблюдений, серединой которых является медиана. Очень часто для графической иллюстрации вариативности признака используется диаграмма «ящики-усы» (Box&Whisker Plot). Внутри «ящика» указывается значение медианы, его размер определяется величиной квартильного размаха, а длина «усов» – максимальным и минимальным значениями показателя в группе испытуемых (рис.6.23).



**Рис.6.23. Иллюстрация вариативности признака для данных из табл. 6.14**

**Дисперсия и стандартное отклонение** – наиболее часто используемые меры изменчивости переменной. Если при вычислении размаха и квартильного отклонения (как и при вычислении моды и медианы) не учитывается каждое отдельное значение переменной, то при вычислении дисперсии и стандартного отклонения (как и

при вычислении среднего) используется каждое значение переменной. Оценка дисперсии совокупности  $n$  значений  $x_1, x_2, \dots, x_n$  со средним  $\bar{x}$  определяется как

$$s_x^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} \quad (6.10)$$

Величина, представляющая собой положительное значение квадратного корня из дисперсии, называется стандартным отклонением или средним квадратичным отклонением:

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}} \quad (6.11)$$

Оценку стандартного отклонения принято обозначать буквой  $s$ . Однако традиционно сложилось так, что в отчётах о психологических исследованиях эта величина обозначается греческой буквой  $\sigma$  (сигма), а не  $s$ . На самом деле,  $\sigma$  – это стандартное отклонение в генеральной совокупности, а  $s$  – оценка этого параметра в исследованной выборке.

Стандартное отклонение считается более удобной характеристикой вариативности признака, чем дисперсия, т.к. она измеряется в тех же единицах, что и исходные значения признака. Теоретически значение стандартного отклонения может быть любым, и чем оно выше, тем сильнее разбросаны значения переменной относительно среднего значения. При  $\sigma = 0$  изменчивость признака отсутствует, т.е. все значения переменной равны.

Выбор меры изменчивости для описания результатов обследования, как и в случае выбора меры центральной тенденции, зависит от шкалы измерения признака, специфики множества полученных данных и целей исследования.

Для шкалы наименований ни одна из вышеперечисленных мер изменчивости не подходит, и для описания вариативности признака используют число классификационных ячеек (например, если в

одной международной конференции участвовали представители пяти государств, а в другой – представители 107 государств, то в первом случае число классификационных ячеек равно 5, а во втором – 107, что говорит о значительно более высокой вариативности, т.е. представительности второй конференции). Для порядковой шкалы предпочтительным является использование размахов (между максимальным и минимальным значением  $X_{\max} - X_{\min}$ ; 90-м и 10-м процентилями  $P_{90} - P_{10}$ ; третьим и первым квартилями  $Q_3 - Q_1$ ), а стандартное отклонение используется для оценки вариативности признака, измеренного по интервальной шкале.

### Контрольные вопросы и задания

1. В таблицах 6.16 – 6.19 представлены результаты обследования когнитивной сферы 4-х групп испытуемых с помощью теста Равена. Вычислите моду, медиану, среднее значение и квартили множества данных своего варианта. Номер варианта соответствует номеру студента в списке группы (табл. 6.15).

**Таблица 6.15. Варианты заданий.**

Группа	В	С	Д	Е	Сумма
1		1	2	3	4
2	5	6	7	8	9
3		10	11	12	13
4	14	15	16	17	18

**Таблица 6.16. Результаты обследования группы 1.**

№	А	В	С	Д	Е	Сумма
1.	11	12	11	11	8	53
2.	11	12	10	10	8	51
3.	11	12	10	11	7	51
4.	11	12	11	9	6	49
5.	11	8	11	9	8	47
6.	12	11	9	7	6	45
7.	11	11	10	8	5	45
8.	12	11	7	9	5	44
9.	11	9	6	9	6	41
10.	11	8	7	10	2	38
11.	12	10	9	8	7	46

**Таблица 6.17. Результаты обследования группы 2.**

№	A	B	C	D	E	Сумма
1.	12	12	11	9	9	53
2.	10	12	10	10	8	50
3.	12	11	10	10	7	50
4.	11	10	10	11	6	48
5.	12	11	10	10	4	47
6.	10	10	8	10	6	44
7.	12	11	9	8	3	43
8.	11	11	9	9	2	42
9.	11	9	7	9	3	39
10.	10	8	9	7	1	35
11.	10	8	6	8	2	34
12.	9	8	7	5	5	34
13.	12	7	9	5	0	33
14.	12	8	4	6	2	32

**Таблица 6.18. Результаты обследования группы 3.**

№	A	B	C	D	E	Сумма
1.	12	11	10	10	6	49
2.	12	12	10	9	6	49
3.	11	11	7	10	7	46
4.	11	12	10	8	4	45
5.	11	11	10	9	3	44
6.	11	10	10	10	3	44
7.	12	11	10	9	1	43
8.	11	12	10	9	0	42
9.	9	11	10	9	1	40
10.	7	12	8	9	1	37
11.	9	9	8	8	3	37
12.	10	10	7	7	1	35
13.	10	10	6	7	0	33

2. Постройте гистограмму распределения частот для данных своего варианта.

3. Постройте график «ящички с усами»: центр ящичка – медиана, границы ящичка – квартильный размах, усы – размах.

4. Вычислите среднее объединённой выборки, состоящей из трёх подгрупп, если известны числовые характеристики каждой подгруппы (табл. 6.20). Напишите формулу расчёта и произведите вычисления.



**Таблица 6.19. Результаты обследования группы 4.**

№	A	B	C	D	E	Сумма
1.	7	4	3	1	0	15
2.	10	11	8	6	4	39
3.	9	11	8	8	2	38
4.	10	10	8	7	6	41
5.	9	8	6	8	2	33
6.	8	11	6	6	2	33
7.	9	11	9	10	1	40
8.	12	10	7	8	2	39
9.	12	11	11	10	2	46
10.	11	8	6	10	1	36
11.	11	11	8	9	7	46
12.	12	11	9	7	7	46
13.	10	8	4	4	0	26
14.	11	11	9	8	0	39

**Таблица 6.20. Числовые характеристики подгрупп**

№ подгруппы	Среднее значение	Стандартное отклонение	Объём подгруппы
1	44	8,6	25
2	52	7,9	7
3	50	7,2	20

5. Заполните таблицу 6.21.

**Таблица 6.21. Меры центральной тенденции и меры изменчивости для данных, измеренных в различных шкалах**

Шкала	Меры центральной тенденции	Меры изменчивости
Номинальная		
Порядковая		
Интервальная		

## 6.4. Распределение признака

Распределением признака называется закономерность встречаемости разных его значений.

Так, при равномерном распределении все значения признака равновероятны.

Нормальное распределение, также называемое гауссовским распределением или распределением Гаусса – распределение вероятностей, которое задается функцией плотности распределения:

$$\varphi(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (6.12)$$

где параметр  $\mu$  – среднее значение (математическое ожидание) случайной величины и указывает координату максимума кривой плотности распределения, а  $\sigma^2$  – дисперсия.

Нормальное распределение играет важнейшую роль во многих областях знаний, особенно в статистической физике. Физическая величина, подверженная влиянию значительного числа независимых факторов, могущих вносить с равной погрешностью положительные и отрицательные отклонения, вне зависимости от природы этих случайных факторов, часто подчиняется нормальному распределению, поэтому из всех распределений в природе чаще всего встречается нормальное (отсюда и произошло одно из названий этого распределения вероятностей).

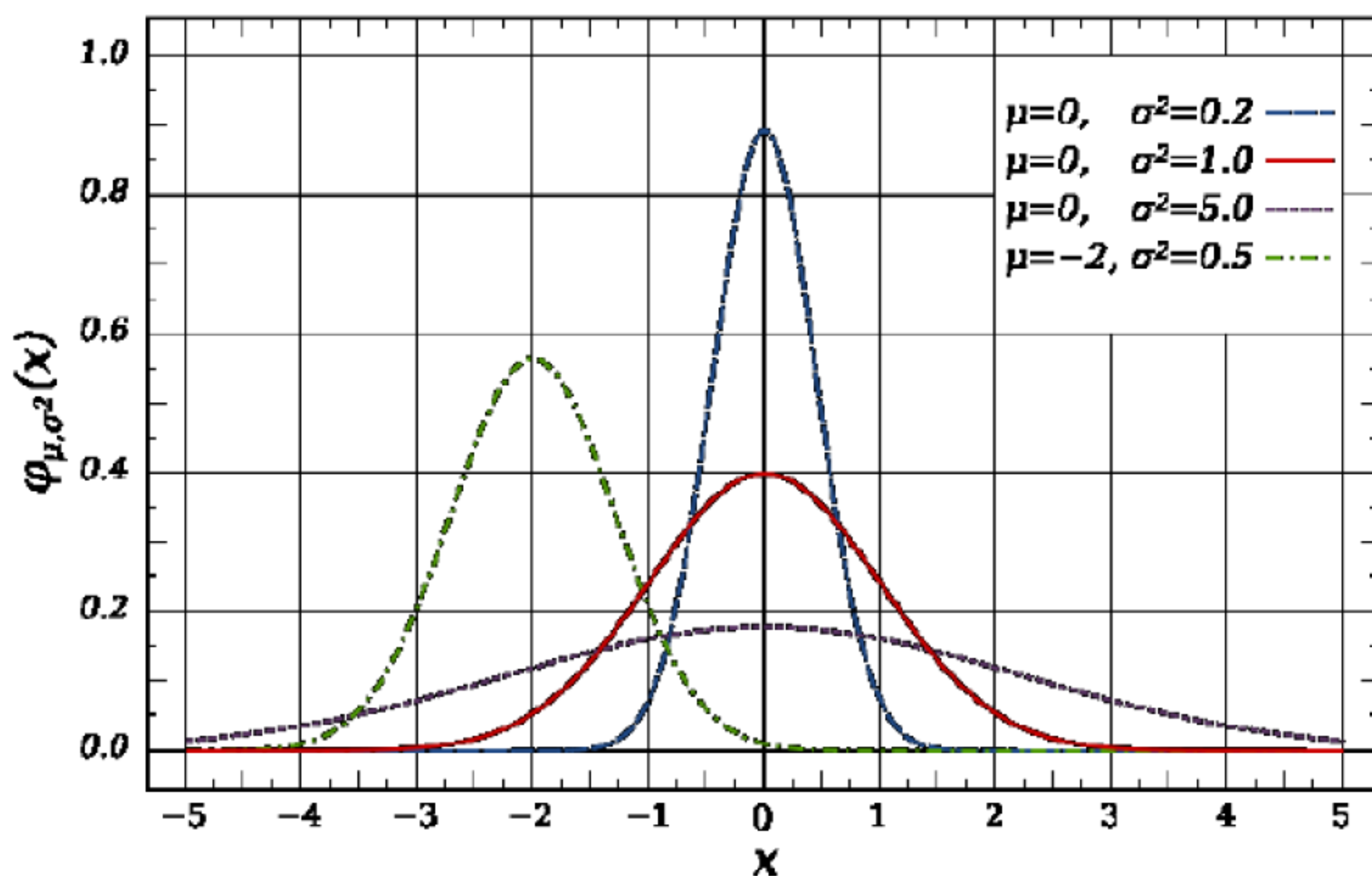
Нормальное распределение зависит от двух параметров – смещения и масштаба, то есть является с математической точки зрения не одним распределением, а целым их семейством. Значения параметров соответствуют значениям среднего (математического ожидания) и разброса (стандартного отклонения).

При нормальном распределении значения, близкие к средней величине, встречаются часто, а положительные и отрицательные отклонения от среднего равновероятны, и частота их появления быстро падает с увеличением отклонения от центра.

Стандартным нормальным распределением называется нормальное распределение с математическим ожиданием  $\mu=0$  и стандартным отклонением  $\sigma=1$

Нормальное распределение очень часто встречается в естественнонаучных исследованиях при изучении переменных, находя-

щихся под воздействием множества факторов. Нормальное распределение было открыто и изучено на примере броуновского движения частицы пыли под воздействием ударов молекул воды, являющегося классической иллюстрацией массового случайного проявления признаков. График нормального распределения представляет собой колоколообразную кривую (рис.6.24).



**Рис.6.24. График нормального распределения при различных значениях  $\mu$  и  $\sigma$**

Нормальное распределение часто встречается в природе. Например, следующие случайные величины хорошо моделируются нормальным распределением:

- отклонение при стрельбе;
- погрешности измерений;
- рост живых организмов.

Психические особенности человека определяются огромным числом факторов (тысячи генов, множество составляющих воспитания, образование, специфика социальной ситуации в различные периоды жизни, психические и соматические заболевания и т.д.)

Следовательно, вполне естественно ожидать нормального распределения тех или иных психических особенностей в популяции людей.

В психологических исследованиях нормальное распределение традиционно применяется при выработке статистических норм психодиагностических методик. Впервые статистические нормы были введены в 20-х годах прошлого века Л.Терменом в ходе стандартизации теста интеллекта Стэнфорд-Бине и с тех пор используются при стандартизации как тестов общих способностей, так и других психодиагностических методик, в частности опросников и проективных методик. Однако, по мнению ряда авторов, нормальное распределение категорий личностной и мотивационной сфер в реальных психологических экспериментах встречается редко. На наш взгляд, «ненормальность» результатов, полученных при обследовании большой репрезентативной выборки испытуемых, в большинстве случаев является следствием несовершенства измерительного инструмента, т.е. психометрических характеристик (в первую очередь, конструктивной валидности) применяемой психодиагностической методики.

Важность оценки характера распределения экспериментальных психодиагностических данных объясняется тем, что многие методы математической обработки требуют нормального распределения исходных данных. Такие методы называются *параметрическими*, т.к. в формулу расчета входят параметры распределения. Если же распределение исходных данных не является нормальным, то следует применять *непараметрические* методы, в основе которых лежит ранжирование либо вычисление процентной доли испытуемых с определёнными характеристиками в общей выборке.

**Параметрами** распределения называются числовые значения, характеризующие

- положение** значений признака на числовой оси (характеризуется мерами положения),
- **изменчивость** значений признака (характеризуется мерами изменчивости),

- **несимметричность** рассеивания значений признака (характеризуется мерой асимметрии),
- **выпуклость** (остроту пика) распределения значений признака (характеризуется мерой эксцесса).

Заметим, что о параметрах распределения можно говорить лишь при проведении обследования всей генеральной совокупности. Как уже отмечалось выше, в реальных психологических экспериментах исследуется выборка, следовательно, по результатам этих экспериментов можно вычислить лишь приближенные значения (выборочные оценки) параметров. С увеличением выборки оценка параметра приближается к его истинному значению. В дальнейшем, говоря о параметрах, будем подразумевать их оценки.

Рассмотрим меры положения, изменчивости, асимметрии и эксцесса нормального распределения.

**Мера положения** – среднее арифметическое  $\bar{x}$  (оценка математического ожидания  $\mu$ ) – вычисляется по формуле (6.13).

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}, \quad (6.13)$$

где  $x_i$  – значения признака ( $i = 1, 2, \dots, n$ ),

$n$  – количество наблюдений.

**Мера изменчивости** – стандартное отклонение  $\sigma$  – вычисляется по формуле (6.14).

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}. \quad (6.14)$$

где  $x_i$  – значения признака ( $i = 1, 2, \dots, n$ ),

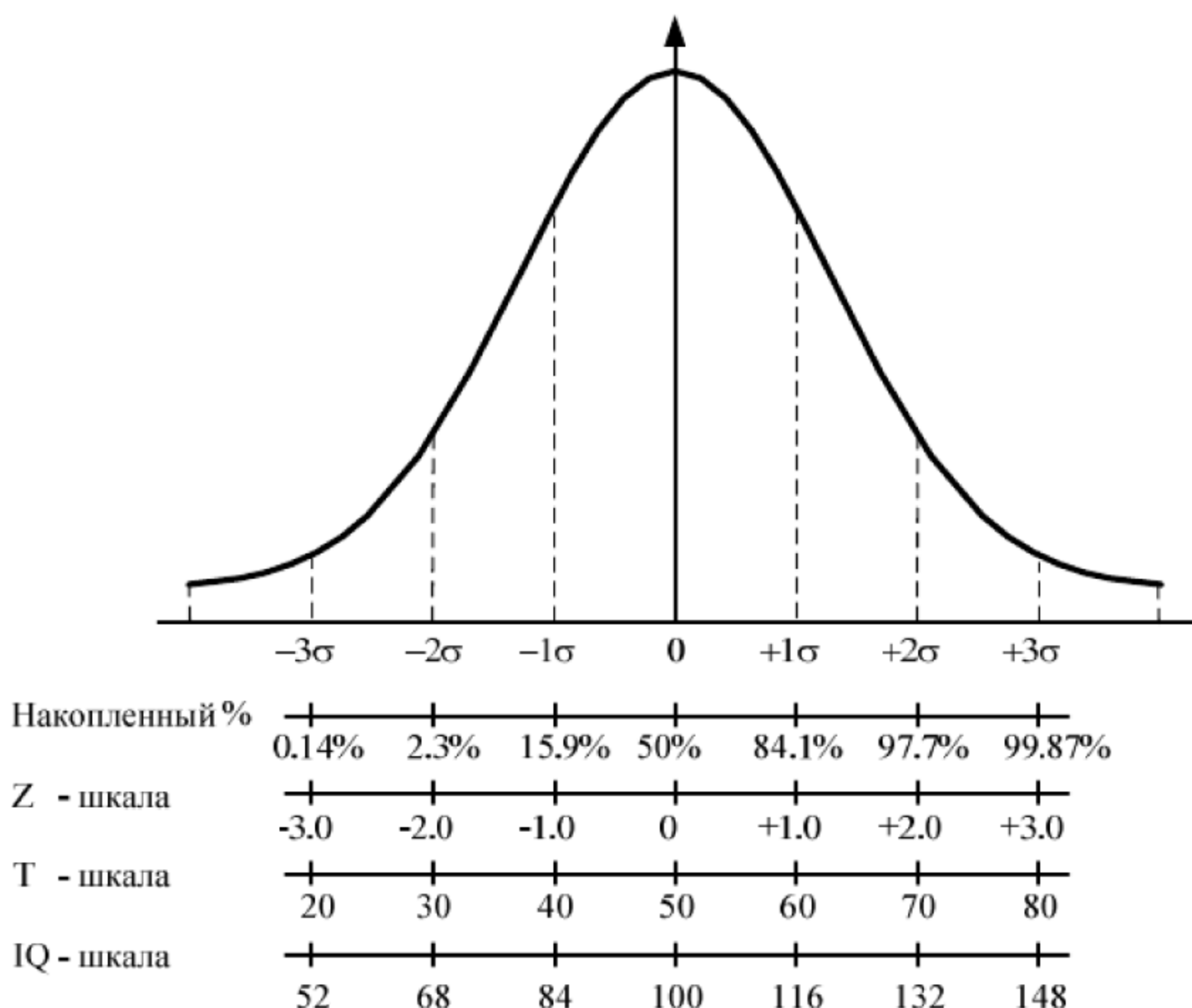
$n$  – количество наблюдений.

Точная форма кривой нормального распределения определяется только этими двумя параметрами ( $\bar{x}$  и  $\sigma$ ). Единичной называется нормальная кривая с  $\bar{x}=0$  и  $\sigma=1$ . Как уже отмечалось, нормальное распределение обладает следующим замечательным свойством:

68% наблюдений лежит в диапазоне  $[\bar{x} - \sigma, \bar{x} + \sigma]$ .

95% – в диапазоне  $[\bar{x} - 2\sigma, \bar{x} + 2\sigma]$ .

99,7% – в диапазоне  $[\bar{x} - 3\sigma, \bar{x} + 3\sigma]$ .



**Рис.6.25. Шкалы стандартных показателей на основе нормального распределения**

Как правило, при выработке статистических норм психодиагностических методик первоначально проверяют нормальность распределения полученных результатов обследования репрезентативной выборки испытуемых.

Если полученные данные нормально распределены, то результаты, лежащие в диапазоне  $[\bar{x} - \sigma, \bar{x} + \sigma]$ , соответствуют норме для данной генеральной совокупности, результаты в интервалах  $[\bar{x} - 2\sigma, \bar{x} - \sigma]$  и  $[\bar{x} + \sigma, \bar{x} + 2\sigma]$ , соответствуют показателям ниже среднего



(выше среднего), а результаты вне диапазона  $[\bar{x} - 2\sigma, \bar{x} + 2\sigma]$  соответствуют значительным отклонениям от нормы в ту или иную сторону (рис. 6.25). Соответственно точки  $\bar{x} + i\sigma$ , где  $i$  принимает целочисленные значения от  $-3$  до  $+3$  определяют уровневые границы значений на шкалах стандартных показателей.

Так, если речь идет о тесте интеллекта, то результаты ниже  $\bar{x} - 2\sigma$ , что соответствует  $IQ < 68$ , могут свидетельствовать об умственной отсталости (олигофрении), а результаты выше  $\bar{x} + 2\sigma$ , что соответствует  $IQ > 132$ , – об интеллектуальной одаренности.

Распределение данных, полученных в результате психологического эксперимента, не всегда бывает строго нормальным. В большинстве случаев общий вид распределения в общих чертах похож на нормальную кривую, но при этом:

- отклонения от среднего в какую-либо сторону встречаются чаще (распределение несимметрично);
- вершина эмпирического распределения выражена слишком сильно либо, наоборот, сглажена.

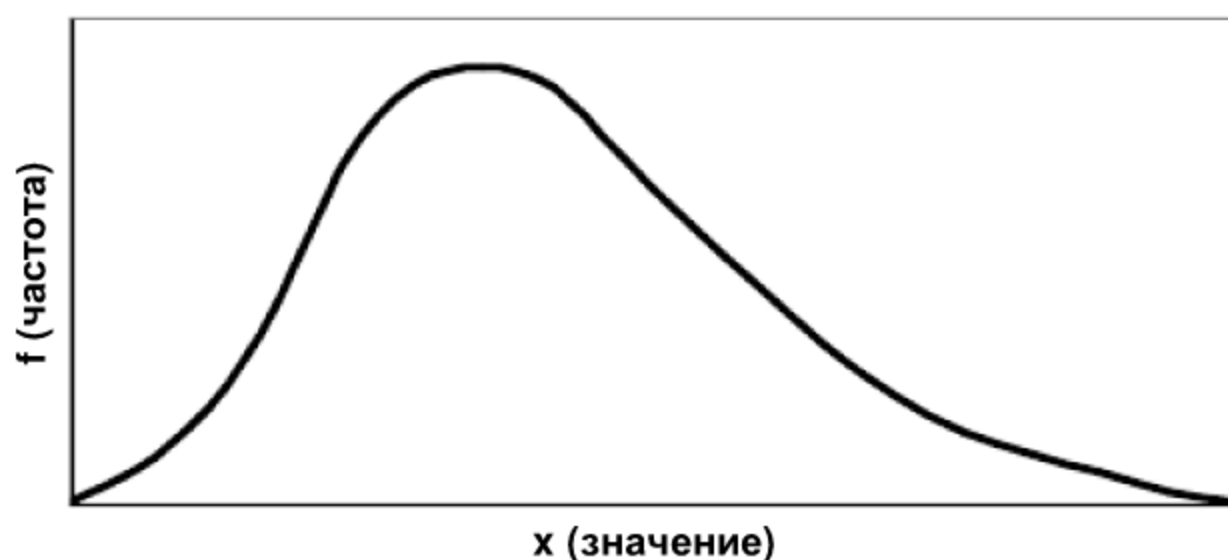
Параметрами распределения, характеризующими эти наиболее часто встречающиеся отклонения, являются меры асимметрии и эксцесса.

Асимметричные распределения получаются при более частом появлении значений, которые больше (или меньше) среднего (повторим, что при нормальном распределении отклонения от среднего в ту или иную сторону равновероятны). При левосторонней, или положительной, асимметрии в распределении чаще встречаются более низкие значения признака, а при правосторонней, или отрицательной – более высокие (рис. 6.26, 6.27).

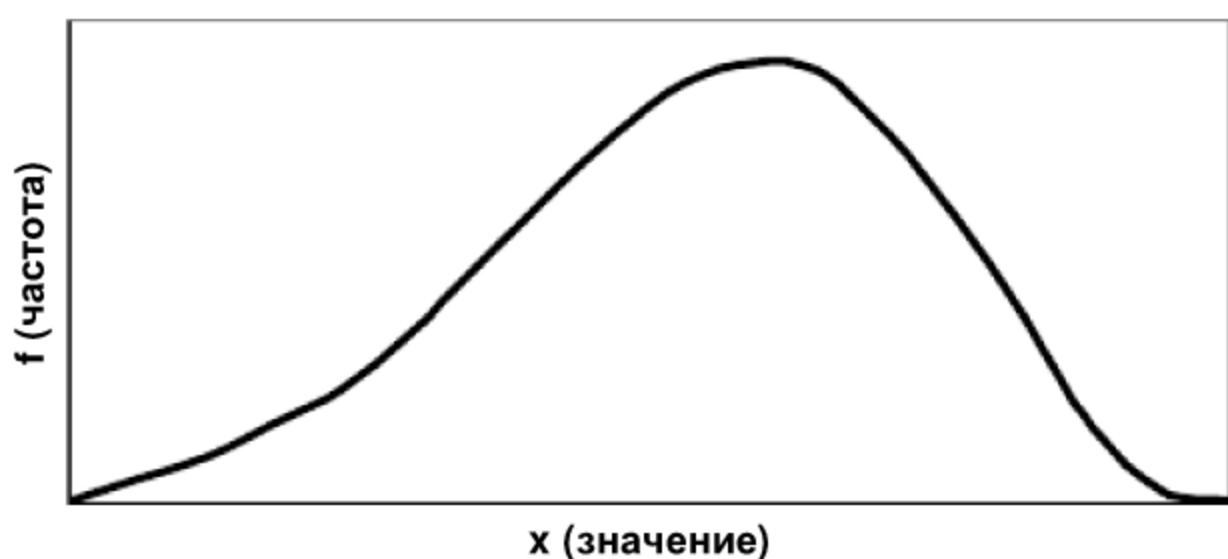
Показатель асимметрии ( $A$ ) вычисляется по формуле (6.15).

$$A = \frac{1}{n\sigma^3} \sum_{i=1}^n (x_i - \bar{x})^3 \quad (6.15)$$

Если распределение симметрично, то  $A=0$ .



**Рис. 6.26. Асимметрия левая, положительная**



**Рис. 6.27. Асимметрия правая, отрицательная**

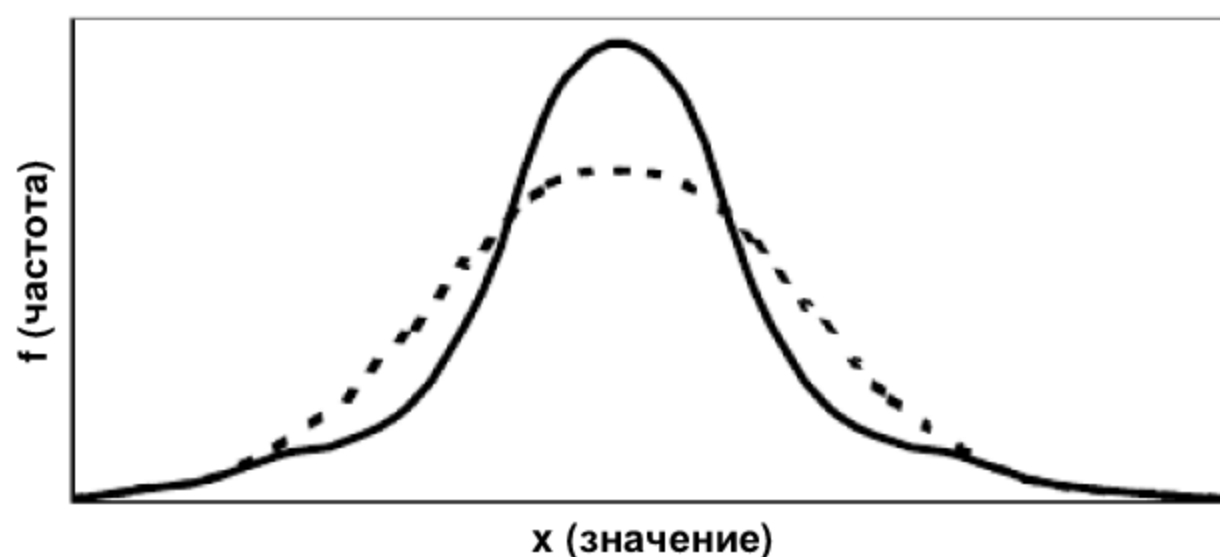
При более частом (по сравнению с нормальным распределением) появлении средних или близких к средним значений образуется распределение с положительным эксцессом (рис. 6.28).

В противном случае распределение характеризуется отрицательным эксцессом и в его центре даже может образоваться впадина, превращающая его в двухвершинное (рис. 6.29).

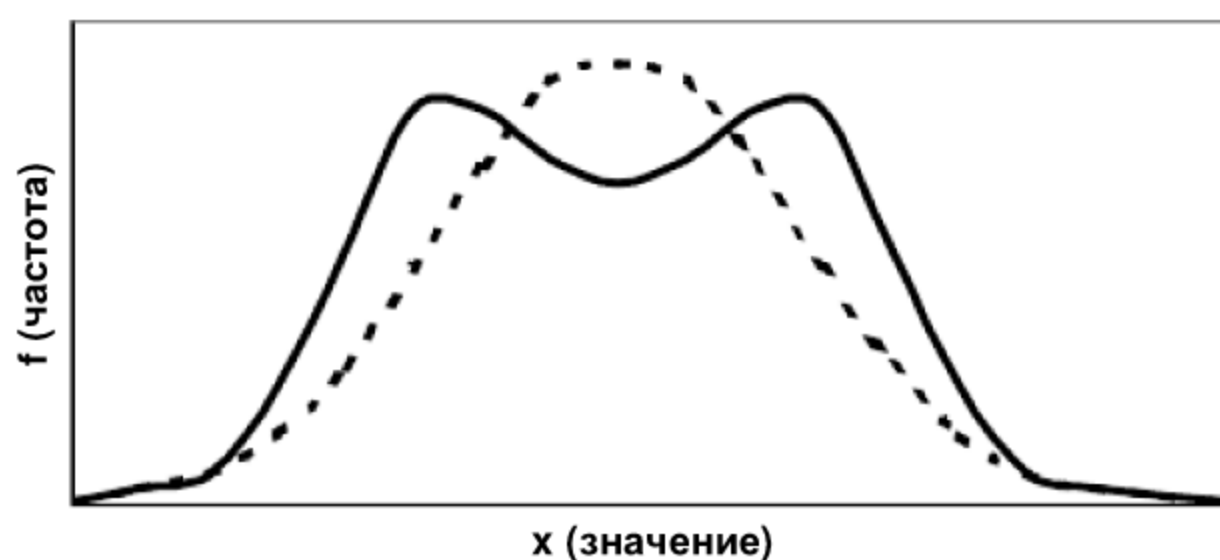
Таким образом, мера эксцесса ( $E$ ) характеризует остроту пика распределения. Она вычисляется по формуле (6.16)

$$E = \frac{1}{n\sigma^4} \sum (x_i - \bar{x})^4 - 3 \quad (6.16)$$

В распределениях с нормальной выпуклостью  $E=0$ .



**Рис. 6.28. Распределение с положительным эксцессом**



**Рис. 6.29. Распределение с отрицательным эксцессом**

Одним из способов проверки распределения экспериментальных данных «на нормальность» является расчет показателей асимметрии и эксцесса и сопоставление их с критическими значениями. Так, показатели асимметрии и эксцесса свидетельствуют о достоверном отличии эмпирических распределений от нормального в том случае, если они превышают по абсолютной величине свою ошибку репрезентативности более чем в 3 раза (метод Н.А.Плохинского):

$$t_A = \frac{|A|}{m_A} \geq 3 \quad (6.17);$$

$$t_E = \frac{|E|}{m_E} \geq 3 \quad (6.18).$$

Показатели асимметрии и эксцесса с их ошибками репрезентативности определяются по формулам 6.19 – 6.22.

$$A = \frac{\sum (x_i - \bar{x})^3}{n\sigma^3} \quad (6.19)$$

$$m_A = \sqrt{\frac{6}{n}} \quad (6.20)$$

$$E = \frac{\sum (x_i - \bar{x})^4}{n\sigma^4} - 3 \quad (6.21)$$

$$m_E = 2\sqrt{\frac{6}{n}} \quad (6.22)$$

Для этих же целей могут использоваться и формулы Е.И.Пустыльника (6.23 – 6.24).

$$A_{кр} = 3 \cdot \sqrt{\frac{6 \cdot (n-1)}{(n-1) \cdot (n-3)}} \quad (6.23)$$

$$E_{кр} = 5 \cdot \sqrt{\frac{24 \cdot n \cdot (n-2) \cdot (n-3)}{(n+1)^2 \cdot (n+3) \cdot (n+5)}} \quad (6.24)$$

Эмпирическое распределение достоверно отличается от нормального распределения, если  $A_{эмп} > A_{кр}$  и  $E_{эмп} > E_{кр}$ .

### Контрольные вопросы и задания

1. Скопируйте рис. 6.24. Постройте на скопированном рисунке графики нормального распределения а)  $\mu=0$ ,  $\sigma^2=3$ ; б)  $\mu=2$ ,  $\sigma^2=0,3$ .
2. Скопируйте рис. 6.25. Добавьте шкалу процентилей и укажите на ней процентиля с шагом 10.

## Глава 7.

### Методы статистического анализа данных

Проводя статистическую обработку результатов психологических измерений важно помнить о том, что методы математической статистики – это очень мощное, но все же лишь дополнение к содержательным методам психологического исследования. **Результаты статистического анализа данных должны являться материалом для дальнейшего глубокого психологического анализа изучаемого феномена.**

Для того, чтобы статистический анализ не представлял собой бессмысленную манипуляцию числами и не привёл к ложным выводам, важно помнить о следующих правилах получения исходных психодиагностических данных:

- измерение должно проводиться с помощью надёжной и валидной психодиагностической методики (либо детально продуманной системы экспертных оценок), выбор которой произведён на основе глубокой теоретической проработки изучаемого психологического явления;
- экспериментальное обследование испытуемых должно быть тщательно спланировано и проведено с соблюдением всех норм регламентации;
- сырые баллы должны переводиться в стандартные показатели (Т-баллы, стандартные баллы) только в том случае, когда обследованная выборка испытуемых принадлежит к той же генеральной совокупности, что и выборка стандартизации.

Перечислим типы задач, встречающихся в исследовательской и практической деятельности психолога в ходе обработки данных психологических обследований:

1. Компактно и информативно представить результаты обследования какой-либо выборки испытуемых. В зависимости от целей исследования подобного рода задачи решаются с помощью методов первичной обработки данных (табулирование, ранжирование, расчёт квантилей), а также с помощью вычисления мер центральной тенденции (мода, медиана, среднее) и мер изменчивости (размах, стандартное отклонение).

2. Сравнить между собой несколько выборок испытуемых для того, чтобы установить, являются ли выборки независимыми или принадлежат одной генеральной совокупности. Здесь используются критерии Стьюдента, Манна-Уитни, Розенбаума.

3. Определить тесноту связи двух рядов показателей, полученных на одной и той же выборке. В такой обработке чаще всего применяют метод корреляций.

4. Определить направление и выраженность сдвигов показателей с течением времени, либо в зависимости от других внешних условий (G-критерий знаков, T-критерий Вилкоксона, критерий Фридмана, L-критерий тенденций Пейджа).

5. Сопоставить два эмпирических распределения или эмпирическое с теоретическим (чаще всего нормальным или равномерным) распределением. Здесь применяются критерий Пирсона и  $\lambda$ -критерий Колмогорова-Смирнова.

6. Изучить зависимость изменений значений зависимой переменной  $Y$  от одной или нескольких независимых переменных  $X$  или предсказать, какими будут значения  $Y$  при определённых значениях  $X$ . В данном случае применяют регрессионный анализ.

7. Изучить влияние отдельных переменных, а также их сочетаний на изменчивость изучаемого признака. Решается с помощью дисперсионного анализа.

8. Провести классификацию испытуемых по результатам измерения нескольких показателей. Для этого используются многомерные статистические методы: дискриминантный анализ и кластерный анализ.



9. Определить структуру взаимосвязей между переменными, сократить их число и перейти от совокупности непосредственно измеряемых переменных к комплексным обобщённым факторам, за которыми стоят комбинации этих исходных переменных. Такого рода проблемы решаются с помощью факторного анализа.

В настоящей главе 7 будут рассмотрены 1–5 типы задач. Обзор методов многомерного анализа данных представлен в главе 8.

## 7.1. Основные понятия статистической обработки данных

### 7.1.1 Статистические гипотезы

Прежде, чем дать определение понятию «статистическая гипотеза», рассмотрим основные этапы научного исследования:

- Выдвижение **научной гипотезы** на основе эмпирического опыта или теории.
- Планирование эксперимента.
- Сбор экспериментальных данных.
- Первичный анализ экспериментальных данных, выдвижение **статистических гипотез**.
- Статистический анализ данных, выводы о принятии или отвержении статистических гипотез.
- Психологический анализ результатов эксперимента и обработки данных.
- Вывод о принятии научной (экспериментальной) гипотезы.

**Научная гипотеза** является основополагающим фактором методологии всего научного исследования, а **статистические гипотезы** (их может быть большое количество) являются рабочим инструментом статистического анализа полученных в ходе исследования данных. Истинность научной гипотезы подтверждается (или опровергается) результатами всего эксперимента. Истинность же статистических гипотез, как будет показано в настоящей главе, проверяется посредством вычисления статистических критериев. Таким образом, научная гипотеза и статистическая гипотеза – это

два разных понятия, имеющих различный вес в общей структуре научного исследования, и их не следует смешивать. В настоящей главе нас будут интересовать гипотезы статистические: их определение, правила выдвижения и способы проверки.

**Статистическая гипотеза** – это утверждение относительно какого-либо признака. Сущность проверки статистической гипотезы заключается в том, чтобы установить, согласуются ли экспериментальные данные с выдвинутым утверждением. При проверке любой статистической гипотезы всегда допускается риск принятия неправильного решения, вероятность которого определяется а) уровнем статистической значимости, на котором принята гипотеза, б) мощностью статистического критерия.

Статистические гипотезы подразделяются на нулевые  $H_0$  и альтернативные  $H_1$ .  $H_0$  и  $H_1$  – это утверждения, взаимно исключающие друг друга.

$H_0$  – это **гипотеза об отсутствии различий** (либо изменений, корреляций – это зависит от содержания исследования) между сопоставляемыми значениями признаков. Нулевая гипотеза – это то, что, как правило, исследователь хочет опровергнуть, когда перед ним стоит задача доказательства значимости различий.

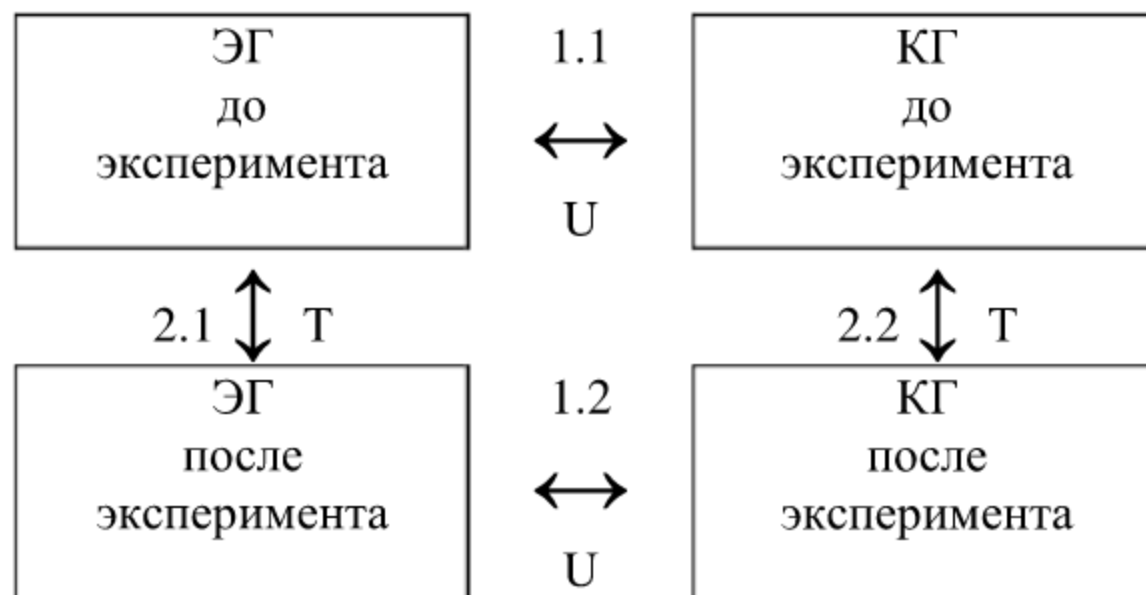
$H_1$  – это **гипотеза о значимости различий**. Как правило, это утверждение, которое доказывается с помощью статистической обработки полученных экспериментальных данных. Но бывают случаи, когда требуется опровергнуть альтернативную гипотезу и принять  $H_0$  (например, доказать отсутствие статистически значимых различий между полученным эмпирическим распределением и нормальным распределением, отсутствие первоначальных различий по значимым характеристикам между контрольной и экспериментальной группами и т.п.).

Нулевая и альтернативная гипотезы бывают направленными и ненаправленными.

**Направленные** гипотезы формулируются в тех случаях, когда нужно проверить значимость и направление различий (в одной группе индивидуальные значения испытуемых по какому-либо

признаку (например, IQ) выше, а в другой ниже, либо если требуется доказать, что в экспериментальной группе произошли более выраженные изменения, чем в контрольной, и в других аналогичных случаях). **Ненаправленные** гипотезы формулируются в случаях, когда нам нужно доказать различие форм распределения признака в группах, при сопоставлении эмпирического распределения признака с теоретическим, при доказательстве отсутствия значимых различий между группами испытуемых, между параллельными формами тестов и т.д.

Рассмотрим пример. Студенты, обучающиеся в одной группе, были произвольным образом разделены на две подгруппы, с одной из которых была проведена серия занятий по развитию памяти. Для проверки эффективности занятий обе подгруппы (экспериментальная – ЭГ и контрольная – КГ) были обследованы до и после проведения тренинга развития памяти. В результате проведённых обследований имеется 4 множества данных, которые схематически изображены на рис. 7.1.



**Рис. 7.1. Два способа обработки экспериментальных данных**

Перед формулировкой статистических гипотез следует провести первичный анализ имеющихся данных (вычислить меры центральной тенденции, меры изменчивости множеств данных, построить диаграммы). Допустим, что по результатам первичного анализа данных было выявлено, что

а) при первом замере показатели в экспериментальной группе «не сильно отличаются» от показателей в контрольной группе;

б) при втором замере по сравнению с первым показатели выросли как в экспериментальной, так и в контрольной группе;

в) при втором замере показатели в экспериментальной группе «в целом выше», чем в контрольной.

Рассмотрим два способа обработки этих данных.

Наиболее естественным и популярным является способ, при котором сопоставляются психодиагностические данные в двух группах сначала до эксперимента (стрелка 1.1 на рис. 7.1), а потом после эксперимента (стрелка 1.2). Какие статистические гипотезы могут быть сформулированы в этом случае?

Во-первых, нам требуется доказать отсутствие различий по уровню развития памяти между контрольной и экспериментальной группами до эксперимента (1.1), в этом случае статистические гипотезы будут ненаправленными:

$H_0$  – уровень памяти у студентов экспериментальной группы **не отличается** от уровня памяти у студентов контрольной группы;

$H_1$  – уровень памяти у студентов экспериментальной группы **отличается** от уровня памяти у студентов контрольной группы.

Во-вторых, в случае отсутствия изначальных различий между группами (т.е. принятия  $H_0$  на предыдущем этапе) для доказательства эффективности занятий необходимо сопоставить данные, полученные после проведения тренинга (1.2), и доказать, что в экспериментальной группе результаты выше. В этом случае статистические гипотезы будут направленными:

$H_0$  – уровень памяти у студентов экспериментальной группы **не превышает** уровень памяти у студентов контрольной группы;

$H_1$  – уровень памяти у студентов экспериментальной группы **превышает** уровень памяти у студентов контрольной группы.

Если с помощью статистических методов (например, с помощью U-критерия Манна-Уитни) будет принята альтернативная гипотеза, то тем самым будет доказана эффективность проведённых занятий: уровень развития памяти в экспериментальной группе действительно стал выше.

При втором способе обработки этих же данных можно попытаться доказать, что изменения, происшедшие в уровне развития памяти у студентов контрольной группы, статистически незначимы, а у студентов экспериментальной группы произошли значимые сдвиги в уровне развития памяти. В этом случае изменится и формулировка статистических гипотез, и метод их проверки. Так, при сопоставлении результатов диагностики экспериментальной группы до и после эксперимента (2.1) гипотезы можно сформулировать следующим образом:

$H_0$  – повышение уровня памяти в экспериментальной группе после тренинга является случайным;

$H_1$  – повышение уровня памяти в экспериментальной группе после тренинга является неслучайным.

При сопоставлении результатов диагностики в контрольной группе (2.2) гипотезы формулируются аналогичным образом:

$H_0$  – повышение уровня памяти в контрольной группе является случайным;

$H_1$  – повышение уровня памяти в экспериментальной группе после тренинга является неслучайным.

Проверку гипотез при этом способе обработки данных можно осуществить, например, с помощью Т-критерия Вилкоксона. Доказательством эффективности тренинга является принятие гипотезы  $H_1$  в первом случае (2.1) при одновременном принятии гипотезы  $H_0$  во втором случае (2.2).

### **7.1.2 Статистические критерии**

Как видно из рассмотренного выше примера, инструментом проверки статистических гипотез являются статистические критерии. По определению Г.В.Суходольского, «статистический критерий – это решающее правило, обеспечивающее надёжное поведение, то есть принятие истинной и отклонение ложной гипотезы с высокой вероятностью» [92, с.291].

В качестве критериев в математической статистике применяют определённые случайные величины (значение критерия), являю-

щиеся функциями изучаемых случайных величин (эмпирических данных, полученных в результате психологического обследования). Иначе говоря, эмпирическое значение критерия вычисляется по определённым правилам (для каждого критерия существуют свои правила вычисления) на основе имеющихся экспериментальных данных.

Полученное эмпирическое значение критерия сопоставляется с критическими значениями, которые определяются по специальным таблицам, исходя из количества наблюдений в исследуемой выборке или числа степеней свободы.

Как правило, для принятия альтернативной гипотезы  $H_1$  необходимо, чтобы эмпирическое значение превышало критическое, хотя для некоторых критериев эмпирическое значение должно быть меньше критического (критерии Манна-Уитни, Вилкоксона и некоторые другие).

Заметим, что одно и то же эмпирическое значение критерия может привести к различным выводам в зависимости от количества наблюдений в исследуемой выборке ( $n$ ) или от числа степеней свободы.

Количество степеней свободы в общем случае определяется как число свободно варьирующих значений в выборке. Так, при выполнении задания назвать любые 10 чисел, сумма которых равна 100, первые 9 чисел могут быть любыми, а 10-е определяется как разность между 100 и суммой этих 9 чисел. Следовательно, в рассмотренном примере число степеней свободы (число свободно варьирующих значений) равно 9, или, в более общем случае,  $n-1$ . Число степеней свободы может ограничиваться также величиной среднего значения, числом градаций фактора и т.д. Определение числа степеней свободы необходимо при использовании критериев Стьюдента, Фишера,  $\chi^2$ , а также в дисперсионном анализе. У каждого статистического критерия существуют свои правила определения числа степеней свободы.

Статистические критерии делятся на параметрические и непараметрические.



**Параметрические критерии** – это критерии, включающие в формулу расчёта параметры распределения (среднее и дисперсию): t-критерий Стьюдента, F-критерий Фишера и др.

**Непараметрические критерии** – это критерии, не включающие в формулу расчёта параметров распределения и основанные на оперировании частотами или рангами: U-критерий Манна-Уитни, T-критерий Вилкоксона, угловое преобразование Фишера и др.

И те, и другие критерии имеют свои преимущества и недостатки. Так, **использование параметрических критериев допустимо лишь тогда, когда значения признака измерены по интервальной шкале** (к которым относятся шкалы стандартных показателей: Z-шкала, T-шкала, стеноны и т.п.) и имеют нормальный закон распределения. Если эти условия выполняются, то использование параметрических критериев является предпочтительным, так как они более мощные.

Непараметрические критерии не требуют нормальности распределения признака; кроме того, шкала измерения может быть порядковой и в некоторых случаях даже номинальной. Вычисление непараметрических критериев гораздо менее трудоёмко, и при небольших выборках данных можно обойтись без компьютера.

При описании статистических методов всегда указывается, какого рода критерий подлежит применению – односторонний или двусторонний. В таблицах уровней значимости значения для одностороннего и для двустороннего критериев даются либо в особых столбцах, либо в таблице указывается, какому значению одностороннего критерия соответствует значение двустороннего. Одно и то же критическое значение, как правило, имеет более низкий уровень статистической значимости для двустороннего критерия.

### **7.1.3 Уровни статистической значимости**

Уровень значимости – это вероятность того, что по результатам вычисления статистического критерия принята гипотеза  $H_1$ , а на самом деле она неверна (т.е. верна нулевая гипотеза  $H_0$ ). Поэтому

при формулировке вывода о принятии гипотезы  $H_1$  необходимо указывать, на каком уровне значимости она принята.

В социальных науках (в том числе в психологии) максимально допустимым для отказа от нулевой гипотезы  $H_0$  уровнем значимости признаётся 5%. Это означает, что при проведении таких же исследований не менее чем в 95% случаях будет получен такой же результат (и, соответственно, не более чем в 5% случаях будет получен противоположный результат). Следовательно, гипотеза  $H_1$  принимается на пятипроцентном уровне значимости, или при  $p \leq 0.05$  (вероятность того, что на самом деле верна  $H_0$ , составляет 5% или 0.05).

Аналогично, когда указывается, что гипотеза  $H_1$  принимается на однопроцентном уровне значимости, или при  $p \leq 0.01$ , то имеется в виду, что вероятность того, что на самом деле верна  $H_0$ , составляет 0.01.

Итак, уровень значимости – это вероятность отклонения нулевой гипотезы, в то время как она верна, т.е. вероятность принятия ошибочного заключения.

**Ошибка, состоящая в том, что по результатам вычисления статистического критерия гипотеза  $H_0$  была отклонена, а на самом деле она верна, называется ошибкой 1 рода.**

Вероятность ошибки 1 рода совпадает с уровнем статистической значимости, на котором принята альтернативную гипотезу. Таким образом, если результаты статистической обработки данных привели нас к принятию альтернативной гипотезы, то необходимо указать уровень статистической значимости такого вывода, т.е. определить вероятность ошибки 1 рода. Например, утверждение «принята гипотеза  $H_1$  ( $p \leq 0.01$ )» говорит о том, что исследователь гарантирует повторение полученного результата в 99% случаев, а вероятность того, что полученные выводы ошибочны, равна, соответственно, 1%.

Исторически сложилось так, что при обработке данных психологических исследований приняты уровни статистической значимости, представленные в табл. 7.1.

**Таблица 7.1 Уровни статистической значимости**

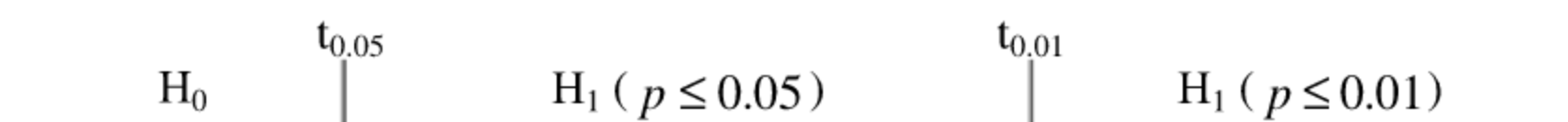
Уровень статистической значимости		Вероятность ошибки 1 рода
достаточный (низший)	$p \leq 0.05$	5%
достоверный	$p \leq 0.01$	1%
высший	$p \leq 0.001$	0,1%

### Правила отклонения $H_0$ и принятия $H_1$

- Если эмпирическое значение критерия равняется критическому значению, соответствующему  $p \leq 0.01$ , или превышает его, то  $H_0$  отклоняется и принимается  $H_1$  на **достоверном** уровне значимости.
- Если эмпирическое значение критерия равняется критическому значению, соответствующему  $p \leq 0.05$ , или превышает его, но не достигает критического значения, соответствующего  $p \leq 0.01$ , то  $H_0$  отклоняется, принимается  $H_1$  на **достаточном** уровне значимости.
- Если же эмпирическое значение критерия ниже, чем критическое значение, соответствующее уровню значимости  $p \leq 0.05$ , то верна нулевая гипотеза  $H_0$ .

Исключения: G-критерий знаков, Т-критерий Вилкоксона, U-критерий Манна-Уитни. Для них устанавливаются обратные соотношения.

На рис.7.2 представлен образец «оси значимости» для t-критерия Стьюдента, иллюстрирующий вышеизложенное правило.



**Рис. 7.2. Пример «оси значимости» для t-критерия Стьюдента**

На этой оси отмечены критические значения  $t_{0.05}$  и  $t_{0.01}$ , соответствующие достаточному ( $p \leq 0.05$ ) и достоверному ( $p \leq 0.01$ ) уровням значимости. Эти два значения делят ось значимости на три области. В зависимости от того, в какую из областей попадёт

эмпирическое значение критерия, полученное в результате вычислений, принимается соответствующее статистическое решение.

#### 7.1.4 Мощность статистического критерия

Мощность критерия – это его способность выявлять различия, если они есть. Иными словами, это его способность отклонить нулевую гипотезу об отсутствии различий, если она неверна.

**Ошибка, состоящая в том, что по результатам вычисления статистического критерия была принята гипотеза  $H_0$ , а на самом деле она неверна, называется ошибкой 2 рода.**

Таким образом, мощность критерия – это его способность не допустить ошибку 2 рода. Статистические критерии отличаются по мощности. Как правило, параметрические критерии более мощные. Непараметрические критерии также имеют разную мощность. Если по результатам вычисления критерия была принята гипотеза  $H_0$ , то следует попробовать применить более мощный критерий, снизив тем самым вероятность ошибки 2 рода. Поэтому для того, чтобы «действовать наверняка», вполне естественно использовать более мощные критерии. Вместе с тем, менее мощные критерии в большинстве случаев имеют более простой алгоритм вычисления, поэтому совсем отказываться от них нецелесообразно, особенно при вычислении вручную.

В таблице 7.2 представлены четыре возможных варианта соотношения статистического решения и фактической оценки  $H_0$ .

**Таблица 7.2 Варианты соотношения статистического решения и фактической оценки  $H_0$**

Статистическое решение	Фактическая оценка $H_0$	
	Верно	Неверно
Принять $H_0$	Правильное решение	Ошибка 2 рода
Отклонить $H_0$	Ошибка 1 рода	Правильное решение

Различия между ошибками 1 и 2 рода иллюстрирует аналогия с вынесением решения суда о виновности или невиновности подсудимого. В соответствии с презумпцией невиновности подсудимый является невиновным ( $H_0$ : подсудимый не совершал преступления) до тех пор, пока не будет доказано противное. Вынося решение, судья может допустить две ошибки: невиновного человека считать преступником (ошибка 1 рода) или же настоящего преступника признать невиновным (ошибка 2 рода). Два других возможных решения, принятых судьёй, являются верными и, следовательно, справедливыми.

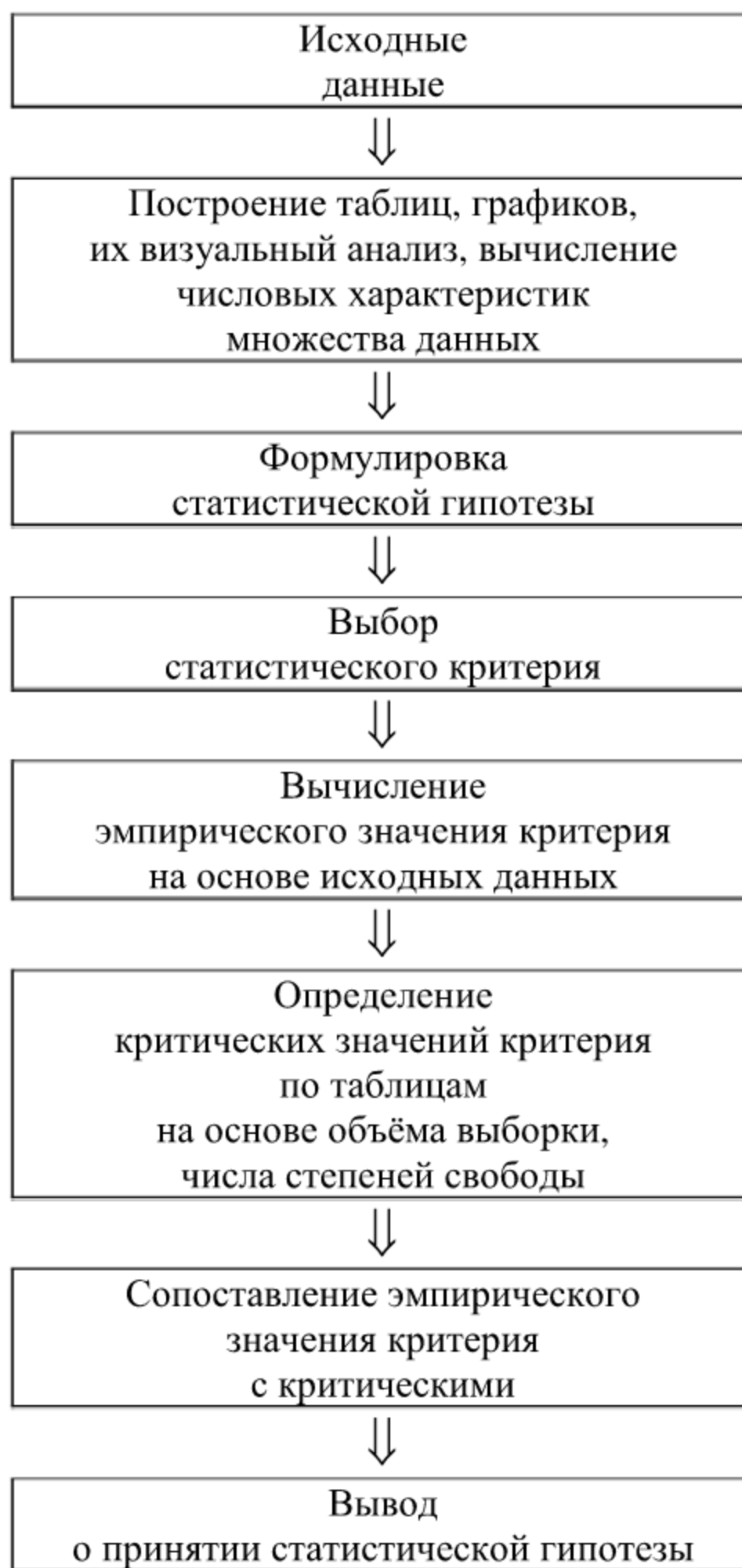
### **7.1.5 Этапы принятия статистического решения**

На начальном этапе статистической обработки исходных данных (численных результатов, полученных по итогам психодиагностического обследования испытуемых) их следует объединить в таблицы, упорядочить, сгруппировать, вычислить меры центральной тенденции и меры изменчивости, оценить тип распределения, сделать графическую иллюстрацию. Эти действия позволят провести первичный анализ данных и наметить план их дальнейшей статистической обработки. Механизм начальной обработки данных подробно описан в главе 6. Следующим шагом (рис.7.3) является формулировка статистических гипотез.

Первичный визуальный анализ таблиц данных обследования, например, двух выборок испытуемых, может привести к предположению о том, что в одной из выборок значения измеренного показателя в целом выше. Это предположение в соответствующей формулировке и является альтернативной гипотезой  $H_1$ , для проверки истинности которой будет применяться статистический критерий.

Выбор статистического критерия, как уже отмечалось выше, зависит от многих обстоятельств: типа задачи, шкалы измерения, распределения признака, количества и связности обследованных выборок, их объёма, а также мощности самого критерия. После того, как критерий выбран, вычисляется его эмпирическое значение

на основе исходных данных. Затем полученное эмпирическое значение сопоставляется с критическими значениями и делается вывод о принятии истинной статистической гипотезы.



**Рис.7.3 Общая схема статистической обработки эмпирических данных**



В заключение отметим, что в зависимости от целей исследования, а также от квалификации и изобретательности исследователя одни и те же психодиагностические данные могут подвергаться статистической обработке различными способами.

### **Контрольные вопросы и задания**

1. *Дайте определения научной гипотезы и статистической гипотезы. На каком этапе научного исследования они выдвигаются? Каким образом проверяется научная гипотеза? статистическая гипотеза?*
2. *Что такое статистический критерий?*
3. *Что такое уровень статистической значимости?*
4. *Сформулируйте определение ошибки 1 рода..*
5. *Изложите правила отклонения  $H_0$  и принятия  $H_1$*
6. *Что такое мощность критерия?*
7. *Сформулируйте определение ошибки 2 рода.*
8. *Перечислите варианты соотношения статистического решения и фактической оценки  $H_0$ ..*
9. *Изложите последовательность этапов принятия статистического решения*

## **7.2. Статистические критерии различий и изменений**

При анализе результатов психологической оценки часто требуется по величине измеренной психологической характеристики сравнить между собой несколько выборок испытуемых для того, чтобы установить, являются ли различия между ними статистически достоверными. При этом обследованные выборки могут быть как несвязными, так и связными (одна и та же выборка испытуемых диагностируется два и более раз в различных условиях, либо с течением времени).

Существует достаточно большое количество критериев различий, каждый из которых имеет свою специфику. В этом параграфе будут изложены правила применения наиболее популярных пара-

метрических и непараметрических критериев для связанных и несвязанных выборок (табл.7.3).

**Таблица 7.3. Статистические критерии различий**

Тип сопоставляемых выборок	Параметрические критерии различий	Непараметрические критерии различий
Связные выборки	t-критерий Стьюдента для связанных выборок F-критерий Фишера	G-критерий знаков, T-критерий Вилкоксона, L-критерий Пейджа, $\chi_r^2$ -критерий Фридмана, $\varphi^*$ -угловое преобразование Фишера.
Несвязные выборки	t-критерий Стьюдента для независимых выборок, F-критерий Фишера	Q-критерий Розенбаума, U-критерий Манна-Уитни, H-критерий Краскела-Уоллиса, S-критерий тенденций Джонкира, $\varphi^*$ -угловое преобразование Фишера.

Для правильного выбора критерия следует учитывать:

- тип измерительной шкалы;
- распределение данных;
- связность сопоставляемых выборок;
- количество сопоставляемых выборок;
- объём выборок (многие критерии имеют ограничения по минимальному и максимальному объёму; некоторые критерии требуют равенства объёма выборок);
- мощность критерия;
- трудоёмкость вычислений.

Начнём рассмотрение критериев различий с непараметрических критериев, которые являются более универсальными, так как правомочность их применения не связана с требованиями использования интервальной шкалы измерения и нормальности распределения признака в обследованной выборке.

### **7.2.1 Непараметрические критерии различий для несвязных выборок**

Использование непараметрических критериев для выявления различий между двумя, тремя и более независимыми выборками испытуемых является одним из наиболее распространённых методов статистической обработки данных. Эти критерии объединяет универсальность и относительная простота применения. Правомочность их использования не связана с необходимостью интервального измерения и нормального распределения данных.

К числу непараметрических критериев различий относятся такие критерии, как Q-критерий Розенбаума, U-критерий Манна-Уитни, H-критерий Краскела-Уоллиса, S-критерий тенденций Джонкира. Первые два критерия используются для сопоставления данных в двух несвязных выборках, остальные – в трёх и более выборках. В настоящем учебнике будет рассмотрен U-критерий Манна-Уитни. Остальные критерии подробно описаны в руководствах [22] и [87].

Критерий Манна-Уитни предназначен для оценки различий между двумя выборками по уровню какого-либо количественно измеренного признака. Выборки могут отличаться по объёму. Этот критерий особенно удобен при малом (до 30-40) объёме выборок, хотя может использоваться и при больших объёмах выборок.

#### **Условия применения U-критерия Манна-Уитни**

- Данные должны быть измерены по порядковой или интервальной шкале.
- Выборки должны быть несвязными. Объёмы выборок могут не совпадать.
- В каждой выборке должно быть не менее 3 наблюдений. Допускается, чтобы в одной выборке было 2 наблюдения, но тогда во второй их должно быть не менее 5.  $n_1, n_2 \geq 3$  или  $n_1 = 2, n_2 \geq 5$ .

- Таблицы критических значений рассчитаны на объёмы выборок  $n_1, n_2 \leq 60$ . Однако при  $n_1, n_2 \geq 30$  вычисления лучше проводить не вручную, а, например, с использованием пакета STATISTICA. Если это невозможно, то вместо критерия Манна-Уитни рекомендуется использовать угловое преобразование Фишера.

#### Порядок вычисления критерия U Манна-Уитни.

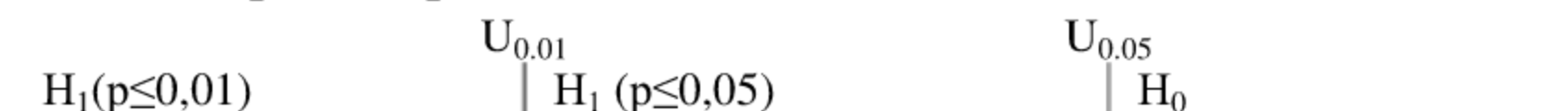
- Объединить выборки и проранжировать результаты, приписывая меньшему значению меньший ранг. Вычислить сумму полученных рангов. Она должна быть равна  $N(N+1)/2$ , где  $N=n_1+n_2$ .
- Посчитать сумму рангов отдельно по каждой выборке (получим  $R_1$  и  $R_2$ ). Проверить, выполняется ли равенство  $R_1 + R_2 = N(N+1)/2$ . Если равенство не выполняется, то при ранжировании и/или суммировании допущена ошибка.
- Определить значение критерия Манна-Уитни по формулам:

$$U_1 = n_1 \cdot n_2 + \frac{n_1(n_1 + 1)}{2} - R_1 \quad (7.1)$$

$$U_2 = n_1 \cdot n_2 + \frac{n_2(n_2 + 1)}{2} - R_2 \quad (7.2)$$

$$U_{\text{эмп}} = \min(U_1, U_2). \quad (7.3)$$

- Определить критические значения  $U_{0,01}$  и  $U_{0,05}$  по таблице и сопоставить с ними полученное эмпирическое значение. Если  $U_{\text{эмп}} > U_{0,05}$ , принимается  $H_0$  (рис. 7.4). Если  $U_{\text{эмп}} \leq U_{0,05}$ ,  $H_0$  отвергается и принимается  $H_1$  ( $p \leq 0,05$ ). Если  $U_{\text{эмп}} \leq U_{0,01}$ ,  $H_0$  отвергается и принимается  $H_1$  ( $p \leq 0,01$ ). Чем меньше значения  $U$ , тем достоверность различий выше.



**Рис. 7.4 Ось значимости для критерия Манна-Уитни.**

- В некоторых таблицах критических значений для каждой пары значений  $n_1$  и  $n_2$  указано не одно, а два числа. В этом случае,

для отклонения  $H_0$  и принятия  $H_1$  необходимо, чтобы меньшее из эмпирических значений критерия было меньше верхней величины, а большее – больше нижней величины.

**Рассмотрим пример.** В исследовании А.Вишняковой экспериментально изучались особенности связи личностных качеств и познавательной мотивации у учащихся 5 класса гимназии г.Обнинска. Требуется установить, существует ли различие в показателях по фактору  $Q_3$  (организованность) опросника СРQ Р.Кеттелла у школьников с высокой и низкой познавательной мотивацией. Исходные данные представлены в табл. 7.4.

**Таблица 7.4. Исходные данные**

Выборка 1 Высокая мотивация			Выборка 2 Низкая мотивация		
№	Испытуе- мый	Q <sub>3</sub>	№	Испытуе- мый	Q <sub>3</sub>
1	А	4	1	К	2
2	Б	2	2	Л	1
3	В	4	3	М	3
4	Г	3	4	Н	3
5	Д	5	5	О	1
6	Е	1	6	П	0
7	Ж	4	Среднее значение		1,67
8	З	4			
9	И	3			
Среднее значение		3,33			

Сопоставление средних значений в выборках школьников с разным уровнем мотивации приводит нас к предположению, что в выборке школьников с высокой познавательной мотивацией оценки по фактору в целом выше.

**Статистические гипотезы:**

$H_0$  – значения по фактору  $Q_3$  у школьников с высокой мотивацией не выше, чем у школьников с низкой мотивацией;

$H_1$  – значения по фактору  $Q_3$  у школьников с высокой мотивацией выше, чем у школьников с низкой мотивацией.

Объёмы выборок 1 и 2 соответственно равны  $n_1=9$ ;  $n_2=6$ . Объединим выборки и проранжируем значения по фактору  $Q_3$ , приписывая меньшему значению меньший ранг. Объём объединённой выборки равен  $N=n_1+n_2=15$ . Упорядоченное множество данных в объединённой выборке представлено в табл. 7.5. Просуммировав значения последнего столбца, убеждаемся, что сумма рангов испытуемых в объединённой выборке совпадает с контрольной суммой  $\sum R = N(N+1)/2=120$ .

**Таблица 7.5. Ранги испытуемых в объединённой выборке**

№	Испытуемый	Мотивация	$Q_3$	Ранг
	Д	высокая	5	15
	А	высокая	4	12,5
	В	высокая	4	12,5
	Ж	высокая	4	12,5
	З	высокая	4	12,5
	Г	высокая	3	8,5
	И	высокая	3	8,5
	М	низкая	3	8,5
	Н	низкая	3	8,5
	Б	высокая	2	5,5
	К	низкая	2	5,5
	Е	высокая	1	3
	Л	низкая	1	3
	О	низкая	1	3
	П	низкая	0	1
Сумма рангов $\sum R$				120

Далее вычислим суммы рангов в выборках с высокой и низкой мотивацией. Для этого разделим объединённую выборку в соответствии с уровнем мотивации испытуемых (табл. 7.6).

Искомые суммы рангов соответственно равны  $R_1 = 90,5$ ;  $R_2 = 29,5$ . Заметим, что  $R_1 + R_2 = 120 = N(N+1)/2$ , что является подтверждением правильности ранжирования.



**Таблица 7.6. Ранги испытуемых с высокой и низкой мотивацией**

Исп.	Мотивация	Q <sub>3</sub>	ранг	Исп.	Мотивация	Q <sub>3</sub>	Ранг
Д	высокая	5	15	М	низкая	3	8,5
А	высокая	4	12,5	Н	низкая	3	8,5
В	высокая	4	12,5	К	низкая	2	5,5
Ж	высокая	4	12,5	Л	низкая	1	3
З	высокая	4	12,5	О	низкая	1	3
Г	высокая	3	8,5	П	низкая	0	1
И	высокая	3	8,5	Сумма рангов R <sub>2</sub>			29.5
Б	высокая	2	5,5				
Е	высокая	1	3				
Сумма рангов R <sub>1</sub>			90.5				

Теперь у нас есть все исходные данные для вычисления эмпирического значения критерия Манна-Уитни. В соответствии с формулами критерия

$$U_1 = n_1 \cdot n_2 + \frac{n_1(n_1 + 1)}{2} - R_1 = 9 \cdot 6 + \frac{9 \cdot 10}{2} - 90,5 = 8,5 \quad (7.4)$$

$$U_2 = n_1 \cdot n_2 + \frac{n_2(n_2 + 1)}{2} - R_2 = 9 \cdot 6 + \frac{6 \cdot 7}{2} - 29,5 = 45,5 \quad (7.5)$$

$$U_{\text{эмп}} = \min(U_1, U_2) = 8,5. \quad (7.6)$$

Критические значения для объёмов сопоставляемых выборок  $n_1=9$ ;  $n_2=6$  равны (см. Приложение 2)

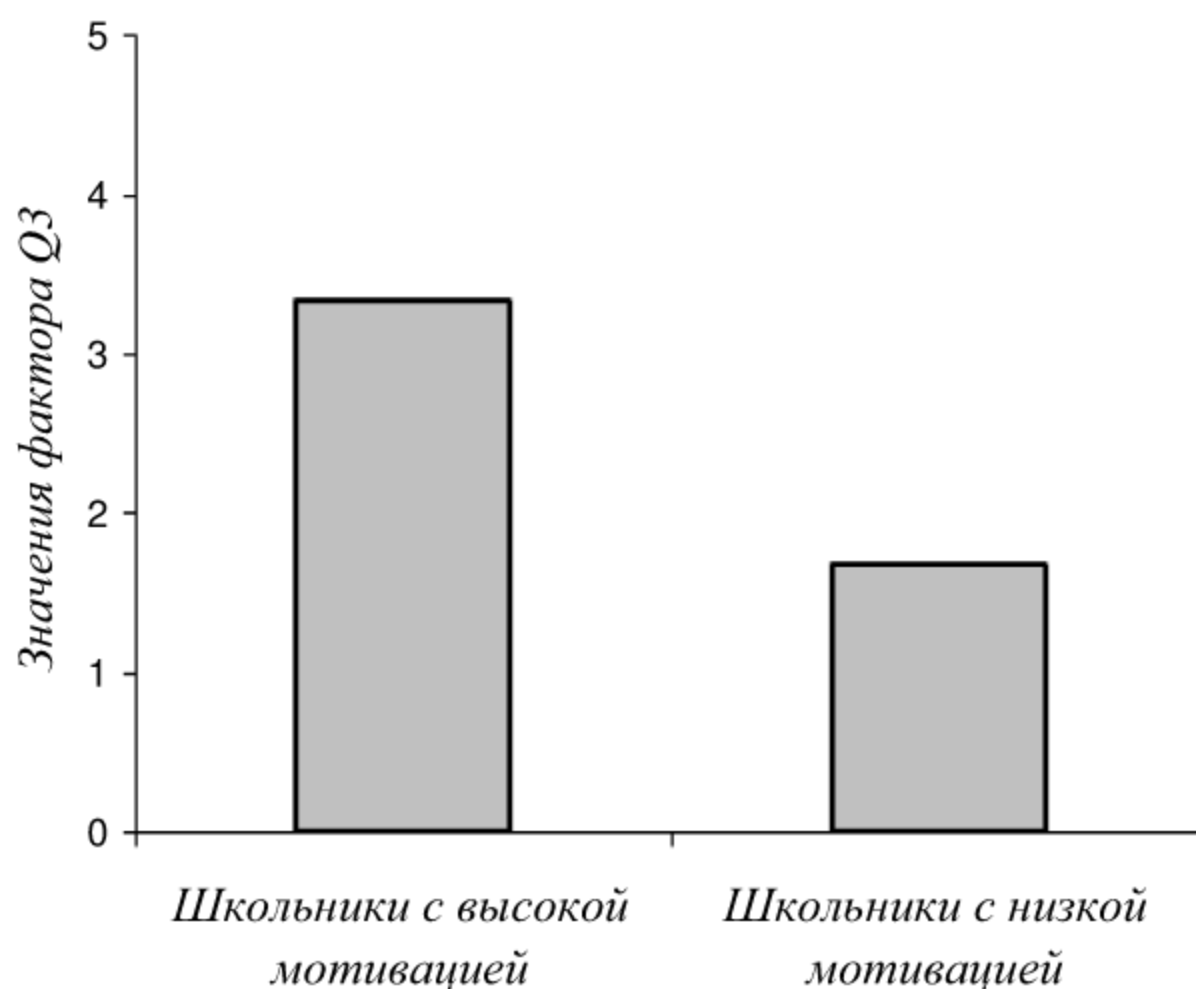
$$U_{\text{кр}}^{0,05} = 12; U_{\text{кр}}^{0,01} = 7 \quad (7.7)$$

Таким образом,  $U_{\text{кр}}^{0,01} \leq U_{\text{эмп}} \leq U_{\text{кр}}^{0,05}$ . Следовательно, отклоняется гипотеза  $H_0$  и принимается гипотеза  $H_1$  на достаточном ( $p \leq 0,05$ ) уровне значимости.

В заключение построим столбиковую диаграмму, иллюстрирующую различие в уровне значений по фактору Q<sub>3</sub>.

Высота столбиков диаграммы (рис.7.5) соответствует среднему значению в каждой выборке (табл.7.4). Ещё раз подчеркнём, построенная диаграмма может служить лишь **иллюстрацией** различий, а **доказательством** их статистической значимости является величина эмпирического значения критерия по сравнению с крити-

ческими значениями для заданных объёмов сопоставляемых выборок.



**Рис. 7.5. Уровень значений фактора  $Q_3$  в выборках школьников с высокой и низкой познавательной мотивацией**

### **7.2.2 Параметрические критерии различий**

Для статистического анализа данных психологических исследований чаще всего применяются два параметрических критерия различий: критерий Стьюдента и критерий Фишера. Первый оценивает различия в средних значениях, а второй – различия между дисперсиями. Серьёзным ограничением правомерности применения этих критериев для обработки результатов психодиагностических экспериментов заключается в требовании к измерению по шкале интервалов и нормальности распределения исходных данных.

#### **Критерий Стьюдента для несвязных выборок**

Критерий Стьюдента используется для сопоставления средних значений в двух выборках данных, которые могут быть не равны по величине. Он позволяет ответить на часто встречающийся при обработке психодиагностических данных вопрос о том, имеют ли статистическую значимость различия в средних значениях двух выборок данных.

Формулы критерия Стьюдента различаются для случаев связанных и несвязанных выборок. В этом разделе будут рассмотрены правила применения критерия Стьюдента для **несвязанных** выборок.

### Условия применения t-критерия Стьюдента

- измерение должно быть проведено в шкале интервалов;
- данные в выборках должны быть распределены по нормальному закону;
- Допускается разное число испытуемых в выборках.

### Вычисление t-критерия Стьюдента

Для вычисления t-критерия Стьюдента существует несколько различных формул. Так, в случае равночисленных выборок ( $n_1 = n_2 = n$ )

$$t = \frac{|\bar{X}_1 - \bar{X}_2|}{\sqrt{m_1^2 + m_2^2}}, \quad (7.8)$$

где

$$m^2 = \frac{\sigma^2}{n}, \quad (7.9)$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}, \quad (7.10)$$

$\sigma$  – стандартное отклонение,  $n$  – объём каждой выборки.

В случае, когда объёмы сопоставляемых выборок не равны ( $n_1 \neq n_2$ ), вычисление критерия производится по формулам:

$$t = \frac{|\bar{X}_1 - \bar{X}_2|}{\sqrt{m_1^2 + m_2^2}}, \quad (7.11)$$

$$m_1^2 + m_2^2 = \frac{\sum (x_{1i} - \bar{x}_1)^2 + \sum (x_{2i} - \bar{x}_2)^2}{n_1 + n_2 - 2} \cdot \frac{n_1 + n_2}{n_1 n_2} \quad (7.12)$$

После вычисления эмпирического значения  $t$  по таблицам критических значений определяется критическое значение  $t_{кр}$  для различных уровней статистической значимости. Вычисление числа степеней свободы осуществляется по формуле

$$k = (n_1 - 1) + (n_2 - 1) = n_1 + n_2 - 2. \quad (7.13)$$

Вывод о принятии статистической гипотезы делается в соответствии с общими правилами. Ось значимости для критерия Стьюдента представлена на рис.7.2.

### **Критерий Фишера**

Как уже отмечалось в главе 6, любое множество данных, полученных в результате обследования выборки испытуемых, можно охарактеризовать с помощью двух показателей, первый из которых является мерой центральной тенденции и определяет типичное значение для данной выборки, а второй является мерой изменчивости и определяет разброс значений. Следовательно, при доказательстве различий между двумя множествами данных недостаточно доказать наличие или отсутствие статистически значимых различий между средними значениями этих множеств. В психологических исследованиях весьма часто встречаются случаи, когда по результатам некоторого воздействия измеряемые переменные «в среднем» не изменились, но существенно изменился разброс значений переменных. И если исследователь при сопоставлении данных ограничивается лишь критерием Стьюдента, то сделанные им на основе такого статистического анализа психологические выводы почти всегда окажутся неполными, а зачастую и ошибочными.

Для сравнения разброса значений переменной в двух выборках испытуемых применяется критерий Фишера. Он позволяет сопоставить величины выборочных дисперсий двух рядов наблюдений. Этот критерий часто применяется в комплексе с критерием Стьюдента. Только совместное применение этих двух критериев может дать полную информацию о различиях между выборками.

Эмпирическое значение критерия Фишера равно отношению дисперсий распределения данных в выборках  $X$  и  $Y$ :

$$F_{\text{эмп}} = \frac{S_x^2}{S_y^2}, \quad (7.15)$$

где

$$S_x^2 = \frac{\sum_{i=1}^{n_x} (x_i - \bar{x})^2}{n_x - 1}, \quad (7.16)$$

$$S_y^2 = \frac{\sum_{i=1}^{n_y} (y_i - \bar{y})^2}{n_y - 1}. \quad (7.17)$$

При вычислении критерия Фишера принято обозначать  $S_x^2$  бóльшую из дисперсий, поэтому всегда выполняется неравенство

$$S_x^2 \geq S_y^2 \quad (7.18)$$

и, следовательно, эмпирическое значение критерия Фишера

$$F_{\text{эмп}} \geq 1. \quad (7.19)$$

Для определения критических значений необходимо вычислить число степеней свободы

$$df_1 = n_x - 1, \quad df_2 = n_y - 1, \quad (7.20)$$

в соответствии с которыми определяются критические значения. Если эмпирическое значение критерия больше критического или равно ему, то различия в дисперсиях статистически достоверны на соответствующем уровне значимости.

### 7.2.3 Статистические критерии изменений

Статистические критерии изменений используются при обработке данных, полученных при двукратном (многократном) обследовании одной и той же выборки испытуемых. С помощью этих критериев проверяются гипотезы о достоверности изменений в измеренных психологических характеристиках. Рассмотрим несколько примеров.

- Выборка испытуемых обследуется по одним и тем же методикам два или более раз через определённые интервалы времени (от нескольких месяцев до нескольких лет). Такие многократные обследования являются частью лонгитюдных исследований.
- Выборка испытуемых обследуется по одним и тем же методикам до и после экспериментального воздействия (например, при изучении эффективности обучающей программы). При проведении таких исследований важно не только доказать наличие достоверного сдвига в измеренных показателях, но и то, что причиной этого сдвига является именно экспериментальное воздействие. Поэтому необходимо наличие **контрольной** группы (выборки, не отличающейся по значимым характеристикам от экспериментальной группы), которая не подвергается экспериментальному воздействию. При этом остальные условия должны быть одинаковыми, чтобы максимально исключить влияние неучтённых факторов. Возможен и следующий вариант: наличие 2-х и более экспериментальных групп, различающихся по исходным характеристикам, внешним условиям, по степени экспериментального воздействия и т.д.

### **Непараметрические критерии изменений**

В параграфе будут рассмотрены два непараметрических критерия изменений: Т-критерий Вилкоксона и L-критерий тенденций Пейджа.

#### **Т-критерий Вилкоксона**

Критерий применяется для сопоставления рядов показателей, полученных в результате двукратного обследования одной и той же выборки испытуемых. Т-критерий позволяет установить, является ли сдвиг показателей в каком-то одном направлении более интенсивным, чем в другом. Критерий Вилкоксона применим в тех случаях, когда признаки измерены в порядковой или интервальной шкале и имеют достаточно широкий диапазон значений, позволяющий проранжировать разницу между вторым и первым значе-



ниями показателя у испытуемых. В ходе первичного анализа данных следует определить более часто встречающееся направление сдвигов, которое называется **типичным**. Сдвиги противоположного направления называются нетипичными.

### **Условия применения Т-критерия Вилкоксона**

- измерение должно быть произведено по шкале порядка или шкале интервалов;
- выборки должны быть связными и одинаковыми по объёму;
- объём выборок – от 5 до 50 человек, но при  $n \leq 12$  критерий недостаточно мощный.

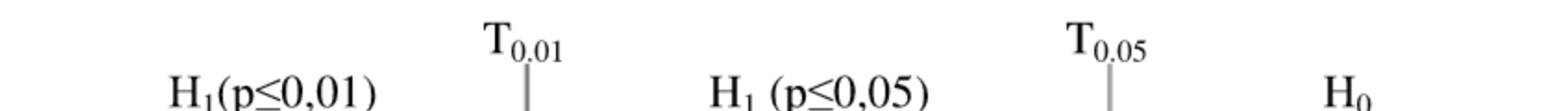
### **Вычисление Т-критерия Вилкоксона.**

- Составить таблицу данных двух замеров на выборке испытуемых.
- Вычислить разность между индивидуальными показателями во втором и первом замерах. Исключить из рассмотрения испытуемых с нулевыми сдвигами; объём выборки при этом уменьшится. Определить, что будет считаться «типичным» сдвигом и сформулировать соответствующие гипотезы.
- Перевести разности в абсолютные величины и записать их отдельным столбцом.
- Проранжировать абсолютные величины разностей, начисляя меньшему значению меньший ранг. Проверить совпадение полученной суммы рангов с расчётной.
- Отметить ранги, соответствующие сдвигам в нетипичном направлении.
- Подсчитать сумму этих рангов по формуле

$$T = \sum R_i, \quad (7.21)$$

где  $R_i$  – ранговые значения сдвигов с более редким знаком.

- Определить критические значения  $T$  для данного  $n$  по таблице критических значений. Если  $T_{\text{эмп}} \leq T_{\text{кр}}$ , то принимается альтернативная гипотеза на соответствующем уровне значимости (рис. 7.6).



**Рис. 7.6. Ось значимости для критерия Вилкоксона**

**Рассмотрим пример.** В исследовании О.А.Коледёнковой (рук. А.В.Хавыло) изучалась динамика работоспособности студентов-психологов в течение рабочей недели. Одним из этапов исследования являлось изучение динамики кратковременной зрительной памяти. Перед началом и в конце учебных занятий в течение 20 секунд студентам демонстрировалась таблица с 12 двузначными числами, которые нужно было запомнить и воспроизвести на бланке. В табл. 7.7 приведены результаты обследования.

**Таблица 7.7 Результаты обследования зрительной памяти студентов**

№	Понедельник		Вторник		Среда		Четверг		Пятница	
	нач.	кон.	нач.	кон.	нач.	кон.	нач.	кон.	нач.	кон.
1	10	2	5	5	6	5	5	3	6	10
2	12	5	8	4	8	5	5	6	7	7
3	6	5	11	10	4	6	7	6	5	7
4	10	3	10	6	7	6	6	6	5	7
5	12	5	12	8	9	5	4	7	6	6
6	10	3	10	5	5	6	7	4	8	8
7	6	4	11	8	4	7	4	8	6	7
8	12	3	12	3	7	7	6	1	9	8
9	9	6	5	6	5	4	5	2	7	2
10	11	4	9	7	3	1	4	2	7	6
11	8	5	5	4	5	5	6	5	6	7
12	10	3	10	5	4	4	3	7	4	4
средн	9,67	4,0	9,0	5,92	5,58	5,08	5,17	4,75	6,33	6,58

Сначала попробуем ответить на вопрос, существуют ли у студентов в начале учебной недели (в понедельник) статистически значимые сдвиги в сторону уменьшения количества верно воспроизведённых чисел.

### Статистические гипотезы:

$H_0$  – сдвиги в сторону уменьшения количества верно воспроизведённых чисел являются случайными;

$H_1$  – сдвиги в сторону уменьшения количества верно воспроизведённых чисел являются неслучайными.

Для проверки гипотез используем Т-критерий Вилкоксона. Для вычисления эмпирического значения критерия вычислим разницу между индивидуальными значениями у каждого испытуемого во вторник до и после занятий и проранжируем абсолютные значения разностей (табл. 7.8).

**Таблица 7.8. Вычисление Т-критерия Вилкоксона**

	До занятий	После занятий	Разность	Ранг разности
1	5	5	0	
2	8	4	4	7
3	11	10	1	2
4	10	6	4	7
5	12	8	4	7
6	10	5	5	9,5
7	11	8	3	5
8	12	3	9	11
9	5	6	-1	2
10	9	7	2	4
11	5	4	1	2
12	10	5	5	9,5
Средн.	9	5,92	$\sum R$	66

Поскольку у испытуемого №1 наблюдается нулевой сдвиг, то его результаты исключаются из рассмотрения, соответственно объём выборки уменьшится:  $N=11$ . Просуммировав значения последнего столбца, убеждаемся, что сумма рангов испытуемых в объединённой выборке совпадает с контрольной суммой:

$$\sum R = N(N+1)/2 = 11(11+1)/2 = 66 \quad (7.22)$$

В табл.7.8 цветом отмечены нетипичные сдвиги (в сторону увеличения числа верно воспроизведённых чисел) и их ранги.

Эмпирическое значение Т-критерия Вилкоксона равно сумме рангов сдвигов в нетипичном направлении (в нашем случае такой сдвиг только один).

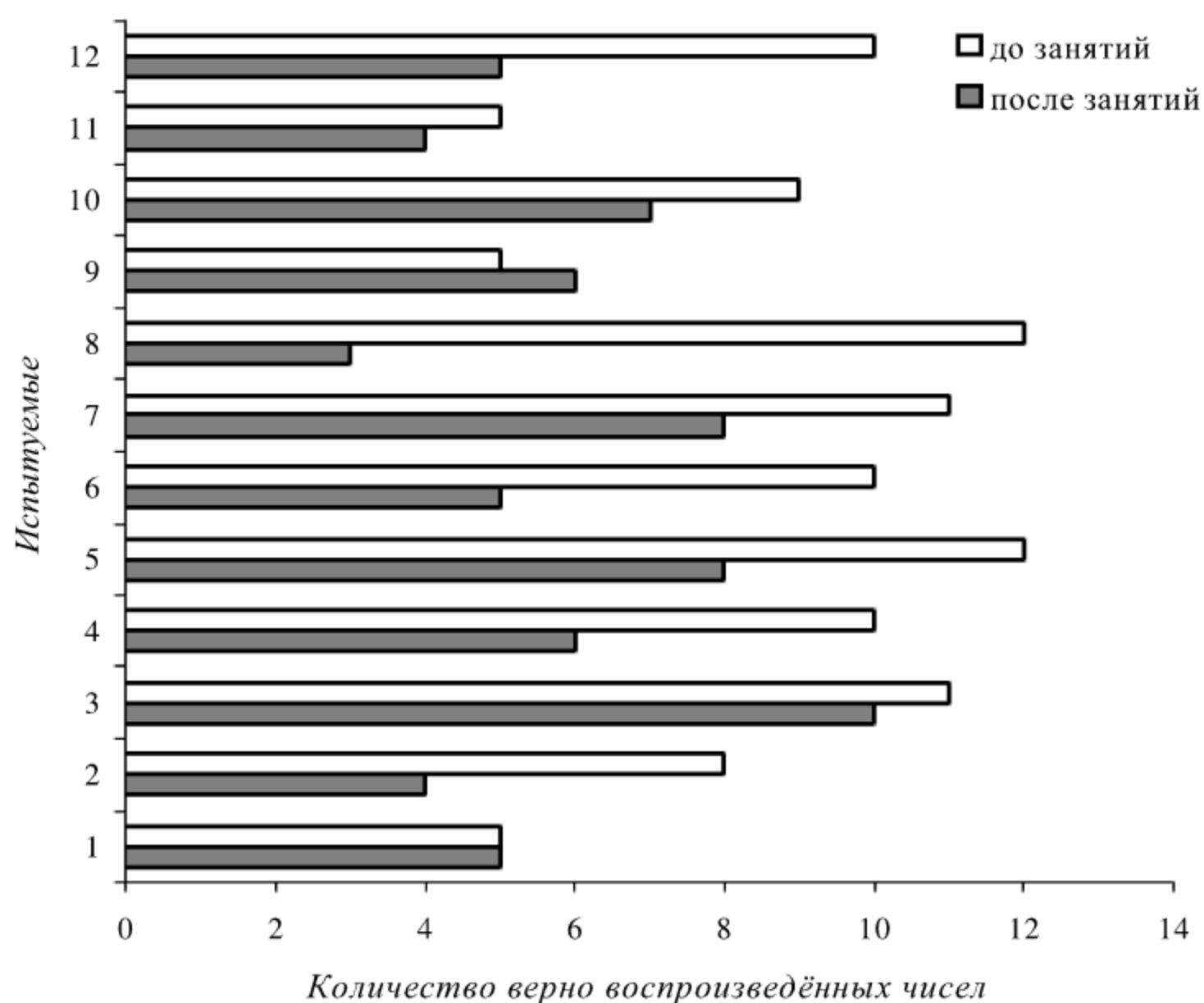
$$T_{\text{эмп}} = \sum R_i = 2. \quad (7.23)$$

Критические значения для объёма выборки  $N=11$  (см. Приложение 2) равны:

$$T_{\text{кр}}^{0,05} = 13; T_{\text{кр}}^{0,01} = 7. \quad (7.24)$$

Сопоставив эмпирическое значение критерия Вилкоксона с критическими, получаем, что  $T_{\text{эмп}} \leq T_{\text{кр}}^{0,01}$ , следовательно, гипотеза  $H_0$  отклоняется и принимается гипотеза  $H_1$  на достоверном ( $p \leq 0,01$ ) уровне значимости. Таким образом, доказано, что во вторник к концу учебного дня у студентов резко снижается объём зрительной памяти. Статистическую значимость изменений зрительной памяти в другие дни недели также можно проверить с помощью критерия Вилкоксона.

В качестве иллюстрации полученного результата можно использовать столбиковую диаграмму (аналогичную изображённой на рис.7.5) или график (аналогичный изображённому на рис.7.9). Но поскольку объём обследованной выборки невелик, в данном случае уместно использование линейчатой диаграммы с индивидуальными результатами каждого испытуемого до и после учебных занятий.



**Рис. 7.7. Изменения в уровне зрительной памяти у студентов в течение учебного дня**

### **L–критерий тенденций Пейджа**

Критерий применяется для сопоставления показателей, измеренных в трёх и более условиях на одной и той же выборке испытуемых. Он позволяет выявить **тенденции** в изменении величин признака при переходе от условия к условию, проверить предположения об определённой возрастной или ситуативно обусловленной динамике тех или иных признаков.

В качестве примера исследования, при обработке данных которого может применяться критерий Пейджа, приведём описание эксперимента С.Милграма [102, с.192].

Одному из двух участников эксперимента предлагается играть роль Учителя (ему объясняется, что целью эксперимента является изучение влияния наказания на память), а второму – роль Учащего-

ся, который получает задание учить пары слов из длинного списка. Учитель должен наказывать Учащегося за каждую совершённую ошибку. В качестве наказания используется электрический разряд, усиливающийся с каждой ошибкой. На самом деле никаких электрических разрядов нет, роль Учащегося играет приглашённый актёр, изображающий страдания от электрического шока, а истинной целью эксперимента является изучение ситуативно обусловленной динамики состояния и поведенческих реакций испытуемого-«Учителя» от «наказания» к «наказанию».

С помощью критерия тенденций Пейджа можно определить степень совпадения индивидуальных тенденций изменения показателей у отдельных испытуемых.

### **Условия применения L-критерия Пейджа**

- шкала измерения порядковая или интервальная;
- выборки связные и одинаковые по объёму;
- число испытуемых не меньше 2 и не больше 12, а количество измерений по каждому – от 3 до 6.

### **Вычисление L-критерия тенденций Пейджа**

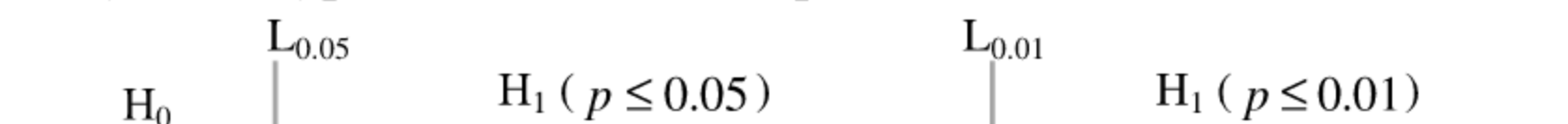
- Проранжировать индивидуальные значения первого испытуемого, полученные им в 1,2,3 и т.д. замерах.
- Прodelать то же самое по отношению ко всем другим испытуемым.
- Просуммировать ранги по условиям, в которых осуществлялись замеры. Проверить совпадение общей суммы рангов с расчётной суммой.
- Расположить все условия в порядке возрастания их ранговых сумм в таблице.
- Определить эмпирическое значение L по формуле

$$L = \sum (T_j \cdot j), \quad (7.25)$$

где  $j$  – порядковый номер, приписанный данному условию в упорядоченной последовательности условий,  $T_j$  – суммы рангов по каждому из условий.



- Определить критические значения  $L_{кр}$  для данного количества испытуемых  $n$  и для данного количества условий  $s$ . Если  $L_{эмп}$  равен критическому значению или превышает его, то утверждение о существовании тенденции принимается на соответствующем уровне значимости (рис. 7.8).



**Рис. 7.8. «Ось значимости» для L-критерия тенденций Пейджа**

Продолжим рассмотрение примера о динамике работоспособности студентов в течение рабочей недели. Одним из этапов исследования являлось изучение динамики кратковременной зрительной памяти студентов. Перед началом учебных занятий в течение 20 секунд студентам демонстрировалась таблица с 12 двузначными числами, которые нужно было запомнить и воспроизвести на бланке. В табл. 7.9 приведены результаты обследования.

**Таблица 7.9. Результаты обследования**

№ испытуемого	Количество верно воспроизведённых чисел				
	Пн	Вт	Ср	Чт	Пт
1	10	5	6	5	6
2	12	8	8	5	7
3	6	11	4	7	5
4	10	10	7	6	5
5	12	12	9	4	6
6	10	10	5	7	8
7	6	11	4	4	6
8	12	12	7	6	9
9	9	5	5	5	7
10	11	9	3	4	7
11	8	5	5	6	6
12	10	10	4	3	4

Попробуем определить, существует ли статистически значимая тенденция изменения количества верно воспроизведённых чисел в течение рабочей недели.

**Статистические гипотезы:**

$H_0$  – не существует общей тенденции изменения индивидуальных результатов в течение учебной недели;

$H_1$  – существует общая тенденция изменения индивидуальных результатов в течение учебной недели.

Для проверки гипотез можно применить критерий тенденций Пейджа, т.к. измерение произведено в порядковой шкале, объём выборки  $n = 12$ , количество условий  $s = 5$ .

Проранжируем индивидуальные результаты каждого испытуемого (табл. 7.10) и просуммируем полученные ранги по дням недели.

**Таблица 7.10. Ранги индивидуальных результатов**

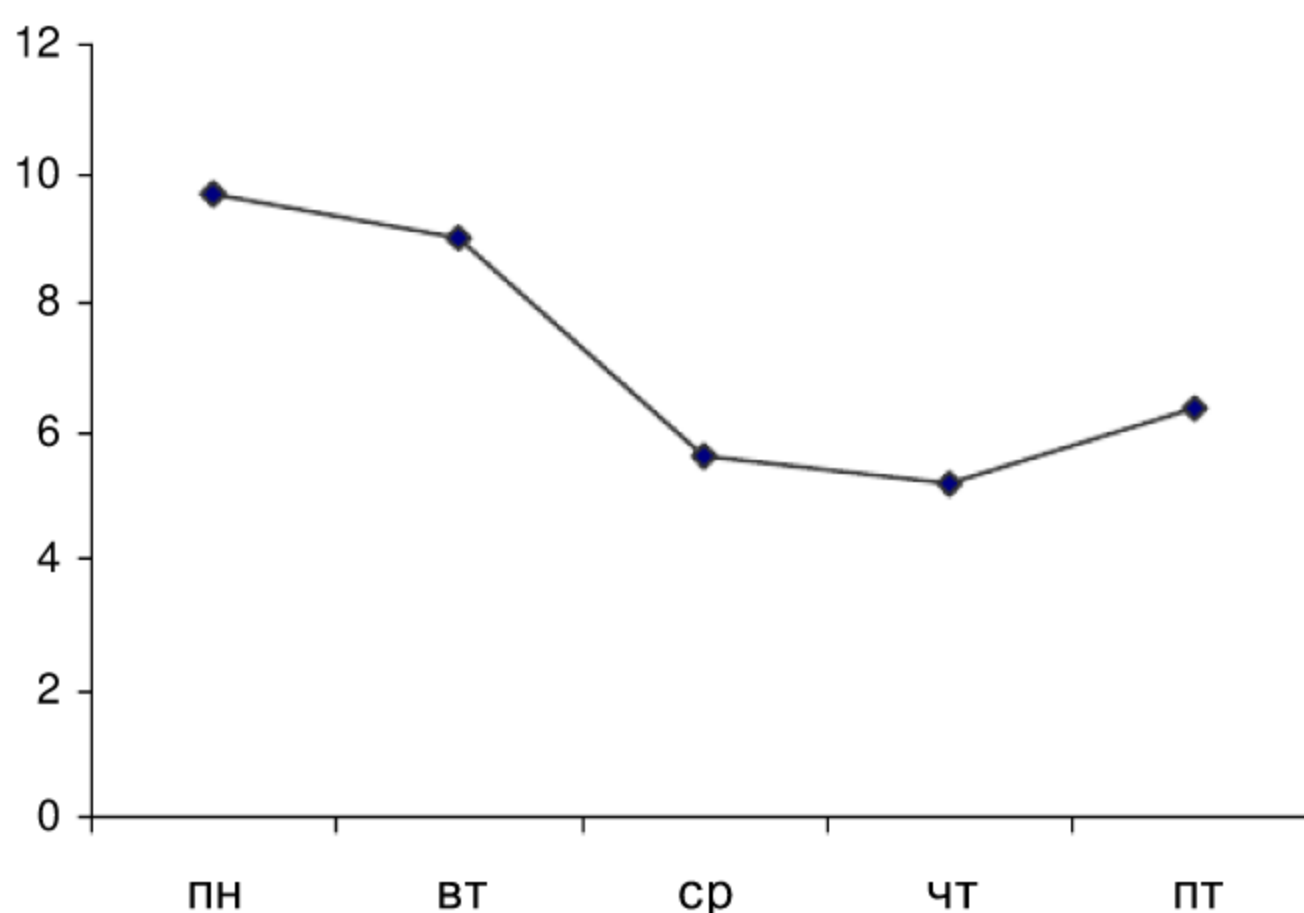
№	Пн		вт		ср		чт		пт	
	рез-т	ранг	рез-т	ранг	рез-т	ранг	рез-т	ранг	рез-т	ранг
1	10	5	5	1,5	6	3,5	5	1,5	6	3,5
2	12	5	8	3,5	8	3,5	5	1	7	2
3	6	3	11	5	4	1	7	4	5	2
4	10	4,5	10	4,5	7	3	6	2	5	1
5	12	4,5	12	4,5	9	3	4	1	6	2
6	10	4,5	10	4,5	5	1	7	2	8	3
7	6	3,5	11	5	4	1,5	4	1,5	6	3,5
8	12	4,5	12	4,5	7	2	6	1	9	3
9	9	5	5	2	5	2	5	2	7	4
10	11	5	9	4	3	1	4	2	7	3
11	8	5	5	1,5	5	1,5	6	3,5	6	3,5
12	10	4,5	10	4,5	4	2,5	3	1	4	2,5
$T_j$		54		45		25,5		22,5		33
$j$		5		4		2		1		3

Из таблицы 7.10 видно, что наименьшая сумма рангов приходится на четверг, затем следуют среда, пятница и вторник. Наконец, в понедельник сумма рангов (соответствующая числу верно воспроизведённых чисел) максимальна.

Эмпирическое значение критерия тенденций Пейджа равно

$$L_{\text{эмп}} = \sum_{j=1}^5 T_j \cdot j = 22,5 \cdot 1 + 25,5 \cdot 2 + 33 \cdot 3 + 45 \cdot 4 + 54 \cdot 5 = 622,5 \quad (7.26)$$

Сопоставив полученное эмпирическое значение с критическими значениями (см. Приложение 2), делаем вывод о принятии гипотезы  $H_1$  на высшем уровне значимости ( $p \leq 0,001$ ). Таким образом, самый высокий уровень развития кратковременной зрительной памяти у студентов-психологов наблюдается в понедельник, затем до четверга наблюдается его снижение, а в пятницу – некоторое повышение. На рис.7.9 представлена графическая иллюстрация динамики уровня кратковременной зрительной памяти у студентов.



**Рис. 7.9. Динамика результатов диагностики кратковременной зрительной памяти у студентов в течение недели**

Анализ графика показывает, что кратковременная зрительная память студентов снижается к середине учебной недели, а к концу недели наблюдается некоторый рост. По результатам данного эксперимента можно рекомендовать составлять расписание занятий с учётом этой тенденции. Это может оказать положительное влияние на эффективность усвоения студентами учебного материала.

#### **7.2.4 Параметрические критерии изменений t-критерий Стьюдента для связанных выборок**

Критерий предназначен для сопоставления двух числовых рядов, полученных на одной выборке (например, до и после экспериментального воздействия).

#### **Условия применения t-критерия Стьюдента для связанных выборок**

- измерение должно быть проведено в шкале интервалов;
- данные в выборках должны быть распределены по нормальному закону;
- число испытуемых в выборках должно быть одинаковым.

#### **Статистические гипотезы:**

$H_0$  – уровень признака в выборке после экспериментального воздействия не превышает уровня признака в выборке до экспериментального воздействия;

$H_1$  – уровень признака в выборке после экспериментального воздействия достоверно превышает уровень признака в выборке до экспериментального воздействия.

Вычисление t-критерия в случае равночисленных связанных выборок ( $n_1 = n_2 = n$ ) производится по формуле (7.27):

$$t = \frac{\bar{d}}{S_d}, \quad (7.27)$$

где

$$\bar{d} = \frac{\sum_{i=1}^n d_i}{n} = \frac{\sum_{i=1}^n (x_i - y_i)}{n}; \quad (7.28)$$

$$S_d = \sqrt{\frac{\sum_{i=1}^n d_i^2 - \frac{(\sum_{i=1}^n d_i)^2}{n}}{n \cdot (n - 1)}}; \quad (7.29)$$

$$d_i = x_i - y_i \quad (7.30)$$

$d_i$  – разность между значениями признака у каждого испытуемого после и до воздействия.  $\bar{d}$  должно получиться положительным,

поэтому вычитать нужно из «большого» ряда «меньший», формулируя гипотезы соответствующим образом.

После вычисления эмпирического значения  $t$  по таблицам критических значений определяется критические значения  $t_{кр}$  для различных уровней статистической значимости. Вычисление числа степеней свободы осуществляется по формуле:

$$k = n - 1 \quad (7.31)$$

В зависимости от полученных результатов делается вывод о принятии статистической гипотезы.

**Пример.** В исследовании Е.Л.Морозовой изучались изменения самочувствия, активности и настроения машинистов электропоездов после рабочего дня. Результаты обследования представлены в табл. 7.11.

**Таблица 7.11 Результаты обследования машинистов до и после рейса**

№	С до	С после	А до	А после	Н до	Н после
1	3,3	3,2	4,9	3,3	7,8	4,4
2	3,1	2,6	2,7	3,1	5,5	5,3
3	5,3	5,1	5,2	3,5	5,4	5,2
4	4,7	4,7	5,5	4,2	4,7	5,4
5	3,1	3,2	5,1	3,5	3,9	4,9
6	3,4	2,1	4,3	3,4	4,1	4,9
7	5,5	3,6	3,5	3,3	5,3	4,3
8	3,9	2,7	3,2	4	4,2	2,7
9	4	4,1	4,1	3,9	4,9	5,1
10	8,7	3,3	3,9	3,3	4,9	4,8
11	6,5	3,4	4	3,2	8,7	4,9
12	5,2	5,2	2,7	5,3	3,9	8,7
13	2,4	5,5	5,3	4,9	3,6	5,1
14	4,2	3,9	3,4	4,3	5,1	4,3
15	2,8	4,2	4,1	4,1	4,8	4,5
Среднее значение	4,4	3,8	5,3	4,9	5,1	5,0

Как видно из таблицы 7.11, среднее значение по каждому из показателей самочувствия, активности и настроения после рейса ни-

же, чем до выхода в рейс. Проверим, является ли такое изменение статистически значимым.

**Статистические гипотезы:**

$H_0$  – уровень показателя самочувствия в выборке машинистов после рейса не ниже уровня показателя в выборке до рейса.

$H_1$  – уровень показателя самочувствия в выборке машинистов после рейса ниже уровня показателя в выборке до рейса. Проверку статистических гипотез произведём с помощью критерия Стьюдента для связанных выборок. Для этого вычислим разности и квадраты разностей индивидуальных показателей самочувствия испытуемых до и после рейса. Результаты вычислений представлены в табл. 7.12.

**Таблица 7.12 Изменения показателей самочувствия машинистов**

№	Самочувствие до рейса	Самочувствие после рейса	d	d <sup>2</sup>
1	3,3	3,2	0,1	0,01
2	3,1	2,6	0,5	0,25
3	5,3	5,1	0,2	0,04
4	4,7	4,7	0	0
5	3,1	3,2	-0,1	0,01
6	3,4	2,1	1,3	1,69
7	5,5	3,6	1,9	3,61
8	3,9	2,7	1,2	1,44
9	4	4,1	-0,1	0,01
10	8,7	3,3	5,4	29,16
11	6,5	3,4	3,1	9,61
12	5,2	5,2	0	0
13	2,4	5,5	-3,1	9,61
14	4,2	3,9	0,3	0,09
15	2,8	4,2	-1,4	1,96

На основе представленных в таблице данных вычислим среднее значение разностей  $\bar{d}$ , а также величину  $S_d$ :

$$\bar{d} = \frac{\sum_{i=1}^n d_i}{n} = \frac{\sum_{i=1}^n (x_i - y_i)}{n} = 0,62; \quad (7.32)$$



$$S_d = \sqrt{\frac{\sum_{i=1}^n d_i^2 - \frac{(\sum_{i=1}^n d_i)^2}{n}}{n \cdot (n-1)}} = 0,496 \quad (7.33)$$

Таким образом, эмпирическое значение критерия Стьюдента равно

$$t = \frac{\bar{d}}{S_d} = 1,249. \quad (7.34)$$

Определим число степеней свободы и критические значения критерия:

$$k = n - 1 = 14; \quad (7.35)$$

$$t_{кр} = \begin{cases} 2.14(p \leq 0.05) \\ 2.98(p \leq 0.01) \end{cases} \quad (7.36)$$

Таким образом, по результатам вычислений принимается гипотеза  $H_0$ , поскольку полученное эмпирическое значение критерия Стьюдента меньше критических.

Аналогичным образом вычисляются эмпирические значения критерия Стьюдента для показателей активности ( $t=1,027$ ) и настроения ( $t=0,29$ ). Сопоставление их с критическими значениями также приводит к принятию нулевой гипотезы.

Таким образом, изменения в средних арифметических показателей самочувствия, активности и настроения у машинистов электропоездов в течение рабочего дня статистически недостоверны. Но для более полного вывода следует сопоставлять не только изменения в средних, но и изменения в вариативности значений признака. Для этих целей применяется критерий Фишера.

### **Комплексное применение критериев Стьюдента и Фишера**

Для иллюстрации эффективности комплексного применения критериев Стьюдента и Фишера продолжим рассмотрение примера об изменении самочувствия, активности и настроения машинистов электропоездов. В предыдущем параграфе было доказано, что показатели самочувствия, активности и настроения машинистов по-

сле возвращения из рейса «в среднем» не изменяются. Однако помимо среднего значения множество данных можно охарактеризовать значением вариативности признака. Посмотрим, как изменилась вариативность значений измеренных показателей.

**Статистические гипотезы:**

$H_0$  – вариация значений показателя самочувствия в выборке машинистов после возвращения из рейса не отличается от вариации этого показателя перед рейсом;

$H_1$  – вариация значений показателя самочувствия в выборке машинистов после возвращения из рейса отличается от вариации этого показателя перед рейсом.

Для проверки гипотез применим F-критерий Фишера. Как уже отмечалось выше, эмпирическое значение критерия Фишера равно отношению дисперсий распределения данных в сопоставляемых выборках X и Y:

$$F_{\text{эмп}} = \frac{S_x^2}{S_y^2}, \quad (7.37)$$

где  $S_x^2 \geq S_y^2$  – значения дисперсий в выборках.

С помощью электронных таблиц Excel вычислим дисперсии значений показателя самочувствия до и после рейса (табл.7.13).

Так как дисперсия показателя самочувствия у машинистов до рейса больше дисперсии этого показателя после рейса, то обозначим

X – значения показателя самочувствия у машинистов до выхода в рейс,  $S_x^2$  – дисперсия;

Y – значения показателя самочувствия у машинистов после рейса,  $S_y^2$  – дисперсия.

Вычислим эмпирическое значение критерия Фишера:

$$F_{\text{эмп}} = \frac{S_x^2}{S_y^2} = \frac{2,73}{1,02} = 2,67. \quad (7.38)$$

и определим число степеней свободы:

$$df_1 = n_x - 1 = 14, df_2 = n_y - 1 = 14 \quad (7.39)$$

**Таблица 7.13. Вычисление дисперсий показателя самочувствия в выборке машинистов до и после рейса**

№	Самочувствие машинистов	
	до выхода в рейс	после рейса
	X	Y
1	3,3	3,2
2	3,1	2,6
3	5,3	5,1
4	4,7	4,7
5	3,1	3,2
6	3,4	2,1
7	5,5	3,6
8	3,9	2,7
9	4	4,1
10	8,7	3,3
11	6,5	3,4
12	5,2	5,2
13	2,4	5,5
14	4,2	3,9
15	2,8	4,2
Среднее	4,41	3,79
Стандартное отклонение S	1,65	1,01
Дисперсия S <sup>2</sup>	2,73	1,02

Критические значения критерия Фишера для числа степеней свободы  $df_{1,2} = 14$  равны (см. Приложение 2)

$$F_{кр} = \begin{cases} 2,48(p \leq 0.05) \\ 3,70(p \leq 0.01) \end{cases} \quad (7.40)$$

Следовательно, принимается гипотеза  $H_1$  на достаточном уровне значимости: происходит стабилизация самочувствия машинистов после рейса (разброс значений значимо уменьшается).

### **Контрольные вопросы и задания**

1. В таблице 7.14 представлены результаты обследования интеллекта мужчин и женщин с помощью теста Равена. Существуют ли статистически значимые различия в уровне интеллектуального развития

- а) женщин и мужчин без учёта уровня образования?
- б) женщин и мужчин со средним образованием?
- в) женщин и мужчин с высшим образованием?
- г) испытуемых с высшим и средним образованием?

#### **Требования к оформлению задач.**

Представить таблицу исходных данных, относящихся к соответствующей задаче. Сформулировать статистические гипотезы. Обосновать выбор критерия. Произвести вычисления. Сделать вывод о принятии гипотезы (в случае принятия гипотезы  $H_1$  указать уровень статистической значимости). Проиллюстрировать результат.

2. Дайте обоснованные ответы на вопросы а) – г). Дайте обоснованные ответы на поставленные вопросы. Исходные данные см. в табл. 7.7.

- а) Существуют ли статистически значимые изменения в уровне памяти студентов в понедельник и пятницу в конце учебного дня?
- б) Существуют ли статистически значимые изменения в уровне памяти студентов в понедельник (в начале и конце учебного дня)?
- в) Существует ли статистически значимая корреляционная связь между результатами тестирования памяти студентов в понедельник в начале и конце учебного дня?
- г) Оцените динамику памяти студентов в течение недели по результатам, полученным в конце учебного дня. Сопоставьте полученный результат с табл. 7.10 и рис. 7.9. Сделайте выводы.

**Требования к оформлению те же, что и в задаче 1.**

**Таблица 7.14. Исходные данные для задания 1.**

№	пол	возраст	Образование.	Результат выполнения теста Равена
1	М	26	С	53
2	М	45	В	54
3	М	22	С	51
4	М	47	С	53
5	М	35	В	52
6	М	59	В	51
7	М	39	С	50
8	М	19	С	50
9	М	45	В	56
10	М	22	С	54
11	М	54	С	47
12	М	36	С	54
13	М	23	С	51
14	М	39	В	48
15	М	36	В	56
16	М	26	В	52
17	М	53	С	49
18	М	56	С	42
19	М	43	С	45
20	М	60	С	42
21	Ж	40	С	42
22	Ж	35	В	54
23	Ж	46	С	52
24	Ж	43	С	49
25	Ж	26	С	43
26	Ж	33	С	47
27	Ж	42	В	44
28	Ж	19	С	52
29	Ж	18	С	57
30	Ж	18	С	49
31	Ж	18	С	49
32	Ж	23	С	46
33	Ж	25	В	55
34	Ж	18	С	50
35	Ж	20	С	49
36	Ж	26	С	49
37	Ж	54	С	48
38	Ж	35	В	44
39	Ж	32	С	55
40	Ж	47	С	56

### 7.3. Статистические критерии согласия распределений

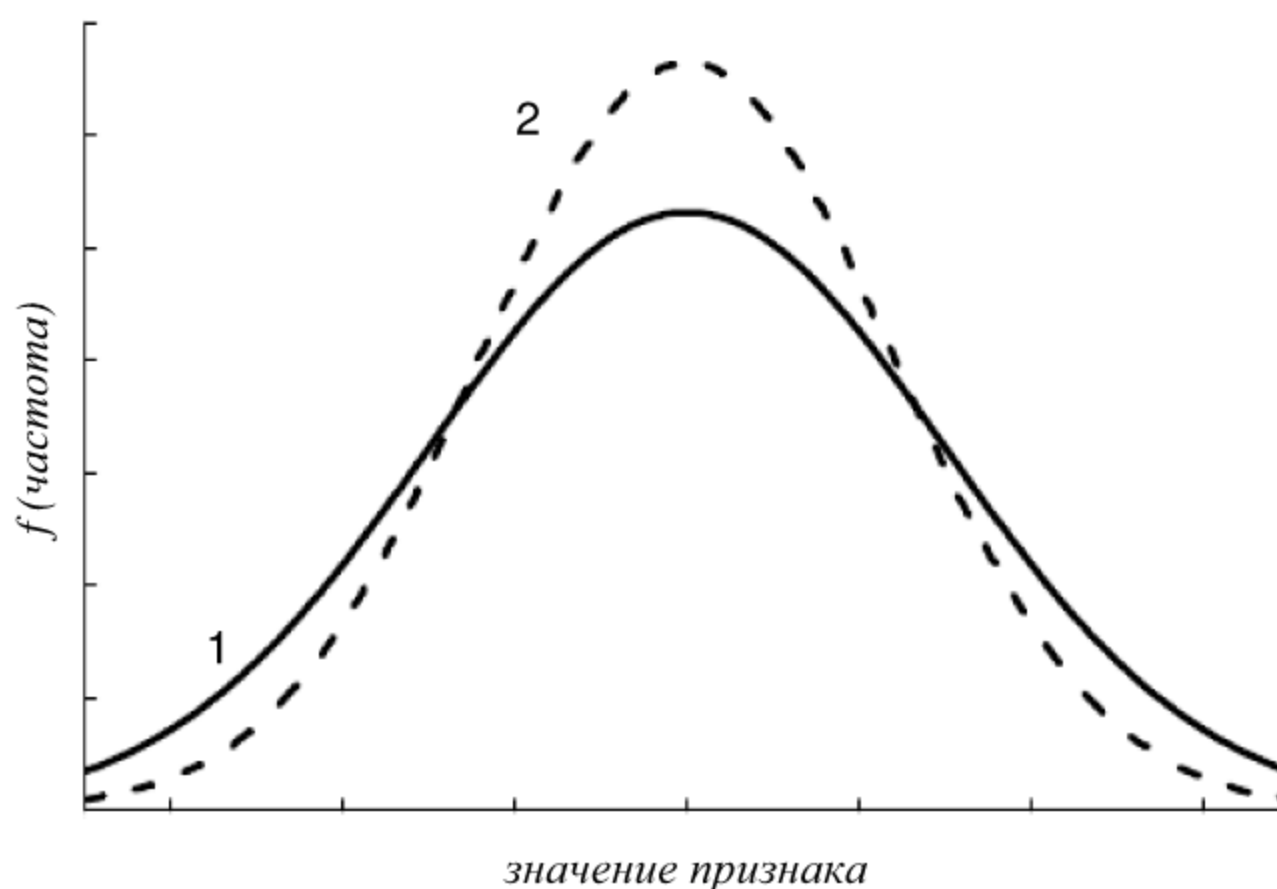
Изучением вероятностных распределений и их свойств математики занимаются уже более 500 лет. Первоначально они использовались при подсчёте шансов в азартных играх, но затем область применения вероятностных распределений значительно расширилась, так как выяснилось, что в разных, на первый взгляд, задачах возникают очень схожие вероятностные законы. В настоящее время для решения задач из различных областей практики применяется множество вероятностных распределений. В психологических исследованиях чаще всего используются нормальное и равномерное распределения, а также распределение  $\chi^2$  при анализе больших массивов данных.

Как уже неоднократно отмечалось, множество статистических критериев, используемых для обработки психодиагностических данных, делится на две группы: параметрические и непараметрические. Параметрические критерии более мощные, и их использование является предпочтительным, но при условии, что данные, подвергаемые обработке, нормально распределены. Следовательно, перед применением любого параметрического критерия необходимо доказать, что статистически значимые различия между эмпирическим и нормальным распределениями отсутствуют. Другим примером задачи сопоставления эмпирического и теоретического распределений является задача доказательства неслучайности предпочтений испытуемых в ситуации свободного выбора. В такого рода задачах полученное эмпирическое распределение данных сопоставляется с равномерным распределением.

Помимо сравнения эмпирического распределения с теоретическим, при обработке психодиагностических данных иногда приходится сопоставлять два и более эмпирических распределений. При отсутствии достоверных различий распределений значения признака в выборках испытуемых можно утверждать, что эти испытуемые относятся к одной генеральной совокупности. В качестве иллюст-



рации (рис. 7.10) приведём пример сравнения распределений фенотипических признаков у мужчин (1) и женщин (2).

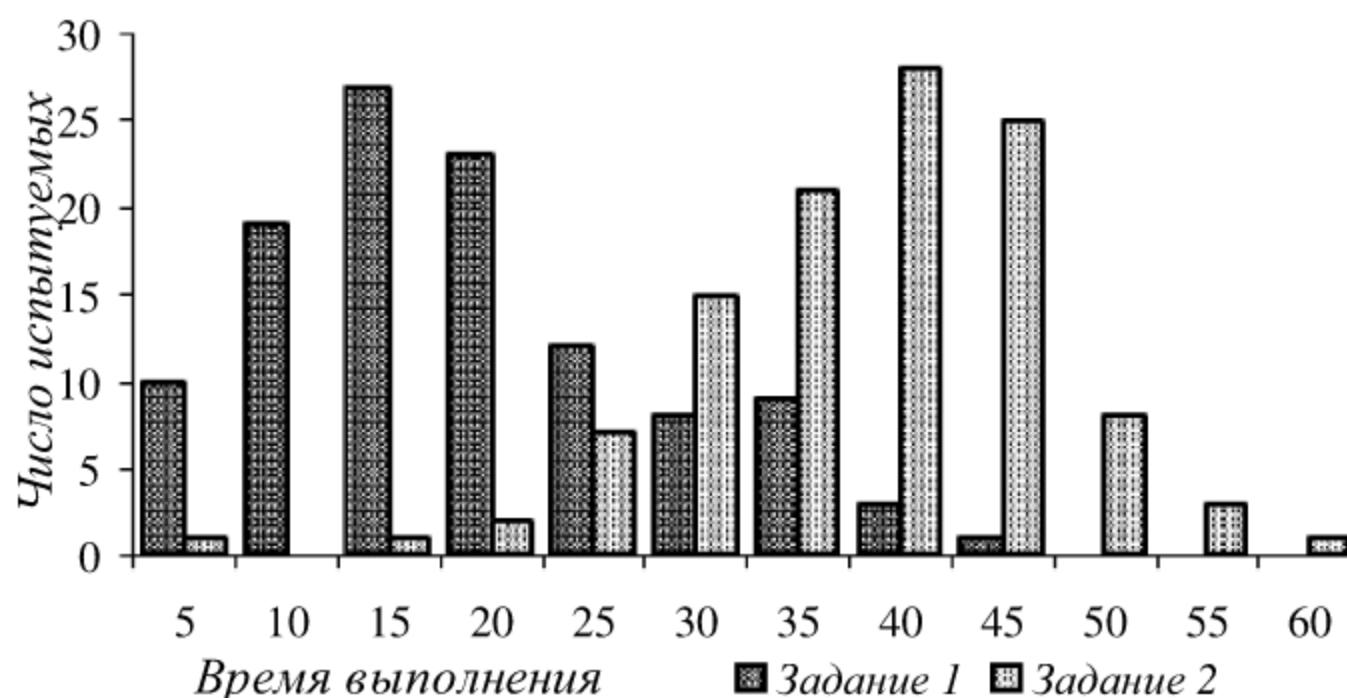


**Рис. 7.10. Распределение фенотипических признаков у мужчин и женщин**

Оба распределения являются нормальными, но в 1 — больший диапазон вариативности и большая дисперсия, чем во 2 распределении. Исследователи связывают это с тем, что мужчины — более авангардная часть популяции, ответственная за поиск новых форм приспособления, поэтому у них чаще встречаются крайние значения (являющиеся как будущими возможными путями эволюции, так и ошибками). Женская же часть популяции ответственна за сохранение уже накопленных изменений, поэтому у них чаще встречаются средние значения фенотипических признаков [87, с.110]. Следовательно, при проведении экспериментальных исследований в группах, смешанных по половому составу, перед сопоставлением результатов в контрольной и экспериментальной группах (см. рис. 7.1) необходимо убедиться в том, что не существует статистически значимых различий в распределении изучаемых показателей в подгруппах мужчин и женщин.

Аналогичные задачи встречаются не только при дифференциации типологий испытуемых, но и при дифференциации тестовых

заданий. Например, для того, чтобы доказать, что два задания имеют различную трудность, можно сопоставить распределения времени выполнения этих заданий (рис.7.11).



**Рис.7.11. Распределение времени выполнения двух тестовых заданий**

Для решения вышеперечисленных задач используются статистические критерии определения расхождения или согласия распределений. Ниже будут рассмотрены критерий  $\chi^2$  К.Пирсона и  $\lambda$ -критерий Колмогорова-Смирнова.

### Критерий $\chi^2$ Пирсона

Критерий  $\chi^2$  (хи квадрат) имеет весьма широкий спектр применения. В настоящем параграфе будет рассмотрена методика использования этого критерия для ответа на вопрос о том, с одинаковой ли частотой встречаются разные значения признака в эмпирическом и теоретическом распределениях или в двух и более эмпирических распределениях.

Ограничения критерия:

- объём выборки должен быть достаточно большим ( $n \geq 30$ ), с ростом  $n$  мощность критерия повышается;

- накопленная частота в каждой ячейке таблицы должна быть не менее 5.

**Порядок вычисления критерия  $\chi^2$  при сопоставлении  
эмпирического распределения с теоретическим**

- Занести в таблицу наименования разрядов и соответствующие им эмпирические частоты  $f_{\text{эмп}}$  (1 столбец).
- Рядом с каждой эмпирической частотой записать теоретическую частоту  $f_{\text{теор}}$  (2 столбец).
- Подсчитать разность между эмпирической и теоретической частотой  $f_{\text{эмп}} - f_{\text{теор}}$  по каждому разряду (строке) – 3 столбец.
- Определить число степеней свободы  $\nu = k - 1$ , где  $k$  – количество разрядов признака. Если  $k=2$ , внести поправку на непрерывность (поправка Йетса), уменьшив разность между эмпирической и теоретической частотами на 0,5:  $f_{\text{эмп}} - f_{\text{теор}} - 0,5$ .
- Возвести в квадрат полученные разности – 4 столбец.
- Разделить полученные квадраты разностей на теоретическую частоту – 5 столбец.
- $\chi^2_{\text{эмп}} =$  сумме значений 5 столбца, т.е.

$$\chi^2_{\text{эмп}} = \sum \frac{(f_{\text{эмп}} - f_{\text{теор}})^2}{f_{\text{теор}}}. \quad (7.41)$$

- Определить критическое значение  $\chi^2$  для данного числа степеней свободы. Если  $\chi^2_{\text{эмп}} < \chi^2_{\text{кр}}$ , то расхождения между распределениями статистически недостоверны. В противном случае расхождения достоверны на соответствующем уровне статистической значимости.

**Порядок вычисления критерия  $\chi^2$  при сопоставлении двух эмпирических распределений.**

- Занести в таблицу наименования разрядов и соответствующие им эмпирические частоты сопоставляемых распределений  $f_{\text{эмп}}^1$  (1 столбец) и  $f_{\text{эмп}}^2$  (2 столбец).
- Вычислить теоретическую частоту  $f_{\text{теор}}$  с учётом пропорции, задаваемой отношением числа наблюдений в выборках (столбцы 3 и 4).
- Подсчитать разности между эмпирической и теоретической частотой  $f_{\text{эмп}} - f_{\text{теор}}$  по каждому разряду в каждой паре столбцов: 1-3 и 2-4. Результат записать соответственно в столбцы 5-6.
- Определить число степеней свободы  $\nu = (c - 1) \cdot (k - 1)$ , где  $c$  – число сопоставляемых распределений, а  $k$  – количество разрядов признака. Если  $\nu = 1$ , внести поправку на непрерывность (поправка Йетса), уменьшив разность между эмпирической и теоретической частотами на 0,5:

$$f_{\text{эмп}} - f_{\text{теор}} - 0,5.$$

- Возвести в квадрат полученные разности – 7-8 столбцы.
- Разделить полученные квадраты разностей на соответствующую теоретическую частоту – 9–10 столбцы.
- $\chi^2_{\text{эмп}} =$  сумме значений 9 и 10 столбцов, т.е.

$$\chi^2_{\text{эмп}} = \sum \frac{(f_{\text{эмп}} - f_{\text{теор}})^2}{f_{\text{теор}}} \quad (7.42)$$

- Определить критическое значение  $\chi^2$  для данного числа степеней свободы. Если  $\chi^2_{\text{эмп}} < \chi^2_{\text{кр}}$ , то расхождения между распределениями статистически недостоверны. В противном слу-

чае расхождения достоверны на соответствующем уровне статистической значимости.

В качестве иллюстрации применения критерия  $\chi^2$  продолжим изучение особенностей динамики работоспособности студентов в течение учебной недели. В ходе эксперимента испытуемые оценивали трудоёмкость учебных дисциплин (табл. 7.15).

**Таблица 7.15. Результаты экспертной оценки трудоёмкости учебных дисциплин**

	1 пара	2 пара	3 пара	4 пара	Всего
Пн	13	10	11	18	52
Вт	3	7	3	4	17
Ср	8	7	7	9	31
Чт	6	13	8	9	36
Пт	5	11	6	10	32

Анализ таблицы 7.15 показывает, что общая учебная нагрузка в среду и в пятницу у студентов примерно одинакова (31 и 32). Вместе с тем (табл. 7.7), в среду показатели работоспособности студентов к концу учебного дня уменьшаются, а в пятницу наблюдается некоторое увеличение. Одной из причин различия в изменениях работоспособности может являться более удачное распределение учебной нагрузки в течение учебного дня в пятницу. Попробуем с помощью критерия  $\chi^2$  установить, существуют ли достоверные различия между распределением учебной нагрузки в среду и в пятницу.

**Статистические гипотезы:**

$H_0$ – распределения учебной нагрузки в среду и пятницу не различаются;

$H_1$ – распределение учебной нагрузки в среду отличается от распределения учебной нагрузки в пятницу.

Вычисление критерия  $\chi^2$  будем производить в соответствии с приведённым выше алгоритмом, записывая результаты вычислений в таблицу 7.16.

**Таблица 7.16. Вычисление критерия  $\chi^2$**

	$f_{\text{эмп}}$		$\Sigma$	$f_{\text{теор}}$		$f_{\text{эмп}} - f_{\text{теор}}$		$\frac{(f_{\text{эмп}} - f_{\text{теор}})^2}{f_{\text{теор}}}$	
	ср	пт		ср	пт	ср	пт	ср	пт
1	2	3	4	5	6	7	8	9	10
1 пара	8	5	13	6,40	6,60	1,60	-1,60	0,40	0,39
2 пара	7	11	18	8,86	9,14	-1,86	1,86	0,39	0,38
3 пара	7	6	13	6,40	6,60	0,60	-0,60	0,06	0,06
4 пара	9	10	19	9,35	9,65	-0,35	0,35	0,01	0,01
$\Sigma$	31	32	63	31	32	0	0	0,86	0,83

Во 2-м и 3-м столбцах таблицы указано распределение общей учебной нагрузки за день по 4-м парам в среду и пятницу. Теоретическая частота нагрузки на каждую пару определяется исходя из суммарной нагрузки на эту пару в среду и пятницу в соответствии с пропорцией, с которой общая нагрузка за эти два дня (63 балла) делится между средой (31 балл) и пятницей (32 балла). Учитывая, что  $13=8+5$  – суммарная нагрузка на первую пару в среду и пятницу, теоретическая частота нагрузки на 1 паре в среду равна  $13 \cdot \frac{31}{63} = 6,4$ , а в пятницу –  $13 \cdot \frac{32}{63} = 6,6$ . Теоретические частоты нагрузки на остальные пары рассчитываются аналогично (столбцы 5 и 6 таблицы). В столбцах 7 и 8 вычисляются разности между соответствующими эмпирическими и теоретическими частотами, а в столбцах 9 и 10 – результат деления  $(f_{\text{эмп}} - f_{\text{теор}})^2$  на  $f_{\text{теор}}$ .

Эмпирическое значение критерия  $\chi^2$  равно сумме элементов столбцов 9 и 10:

$$\chi^2_{\text{эмп}} = 1,7 \quad (7.43)$$

Критические значения критерия  $\chi^2$  для числа степеней свободы  $\nu = k - 1 = 3$  равны (см. Приложение 2)



$$\chi_{\text{кр}}^2 = \begin{cases} 3,84(p \leq 0.05) \\ 6,63(p \leq 0.01) \end{cases}. \quad (7.44)$$

Поскольку полученное эмпирическое значение критерия  $\chi^2$  оказалось меньше критических, статистически значимых различий между распределением учебной нагрузки в среду и пятницу не существует, и причина различий в динамике работоспособности студентов не связана с особенностями расписания занятий.

### **$\lambda$ -критерий Колмогорова-Смирнова**

$\lambda$ -критерий Колмогорова-Смирнова предназначен для сопоставления двух распределений (эмпирического и теоретического или двух эмпирических), а также позволяет найти точку, в которой сумма накопленных расхождений между двумя распределениями является наибольшей, и оценить достоверность этого расхождения.

Если в методе  $\chi^2$  сопоставлялись частоты двух распределений по каждому разряду, то здесь сопоставляются накопленные к данному разряду частоты (т.е. сопоставляются частоты по первому разряду, затем по сумме 1 и 2, затем по сумме 1,2 и 3 разрядов и т.д. Если различия между двумя распределениями существенны, то в какой-то момент разность накопленных частот достигнет критического значения. Порядок вычисления  $\lambda$ -критерия представлен в табл. 7.17.

#### **Ограничения $\lambda$ -критерия:**

- при сопоставлении двух эмпирических распределений необходимо, чтобы объёмы обеих выборок  $n_{1,2} \geq 50$ ;
- при сопоставлении эмпирического распределения с теоретическим объём выборки  $n \geq 5$ ;
- разряды должны быть упорядочены по нарастанию или убыванию какого-либо признака (в противном случае накопление частот будет отражать лишь элемент случайного соседства разрядов).

**Таблица 7.17 Порядок вычисления  $\lambda$  -критерия**

Эмпирическое и теоретическое распределения	Два эмпирических распределения
Занести в таблицу наименование разрядов и соответствующие им эмпирические частоты	То же для двух распределений (2 столбца)
Подсчитать относительные эмпирические частоты (частоты), разделив эмпирическую частоту каждого разряда на общее количество наблюдений	То же для двух распределений (делить на соответствующее число наблюдений)
Подсчитать накопленные эмпирические частоты на каждом разряде	То же для обоих распределений
Подсчитать накопленные теоретические частоты	–
Вычислить разности между накопленными эмпирическими и теоретическими частотами на каждом разряде. Получим $d$	$d$ = разность между накопленными частотами по каждому разряду. $d_{\max}$
Определить максимальное значение $d_{\max}$ и сопоставить его с критическим значением для данного количества наблюдений. Если $d_{\max} \geq d_{\text{кр}}$ , то различия между распределениями достоверны	$\lambda = d_{\max} \sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}}$ . По таблице определить, какому уровню статистической значимости соответствует $\lambda$

**Таблица 7.18. Исходные данные**

Варианты ответа на вопрос анкеты: «Что Вам мешает учиться лучше?»	Факультеты			
	ФЭФ	ФЕН	К	СЭФ
Меня вполне устраивает, как я учусь	25,5%	35,5%	28,3%	36,0%
Неумение организовать своё время, планировать учебную работу	24,5%	29,4%	32,7%	28,2%
Интернет, компьютерные игры, развлечения	25,5%	13,2%	24,4%	17,7%
Я стараюсь, но у меня не получается учиться лучше	16,7%	21,9%	12,7%	12,3%
Работа/подработка	13,1%	9,2%	8,8%	10,5%
Вредные привычки (алкоголь, курение, наркомания)	11,7%	4,8%	8,3%	8,1%
Жилищные условия, при которых невозможно заниматься	9,2%	9,2%	9,3%	7,8%

## Контрольные вопросы и задания

*В таблице 7.18 представлено распределение ответов студентов (1048 чел, несколько вариантов ответа) на вопрос «Что Вам мешает учиться лучше?»*

*С помощью критериев согласия распределений выполните следующие задания:*

*1. Сопоставьте распределение каждого варианта ответов по факультетам с равномерным распределением.*

*2. Сопоставьте распределение ответов между факультетами.*

**Требования к оформлению задач** см. в заданиях к параграфу 7.2

## 7.4. Меры связи

При анализе результатов психологических экспериментов часто возникает вопрос о том, как связаны между собой две переменные в группе обследуемых. Например, имеют ли более высокие оценки по математике в 3-м классе дети, научившиеся читать до 5 лет, по сравнению со своими сверстниками, научившимися читать позже? Или как связаны между собой оценка по показателю оригинальности теста Торренса и число баллов по результатам теста Равена?

Для ответа на приведённые выше вопросы традиционно используются различные коэффициенты корреляции (выбор конкретной меры связи зависит от шкалы измерения, распределения признаков и других аспектов). Хотя зачастую для доказательства связи между значениями двух переменных можно обойтись и без них.

Так, в параграфе 7.2 был рассмотрен пример об изучении особенностей связи личностных качеств и познавательной мотивации у школьников. С помощью критерия Манна-Уитни было доказано, что существует статистически значимое различие в показателях по фактору  $Q_3$  (организованность) опросника Р.Кеттелла у школьников с высокой и низкой познавательной мотивацией. Полученный результат можно считать доказательством связи между организованностью и познавательной мотивацией у подростков.

В общем случае корреляционная связь – это согласованные изменения двух переменных. Если переменных больше двух, то их согласованные изменения называются множественной корреляционной связью. Отметим, что с помощью коэффициентов корреляции определяется именно **связь** между переменными, а не **зависимость** (т.е. **влияние** одной переменной на другую). Хотя в некоторых исследованиях можно с большой долей уверенности говорить именно о зависимости (например, зависимость урожайности от метеорологических условий).

При обработке данных психологических исследований для определения корреляционной связи между переменными чаще всего применяются параметрический коэффициент корреляции Пирсона  $r_{xy}$ , коэффициенты ранговой корреляции Спирмена  $r_s$ , и  $\tau$  Кендалла, коэффициент ассоциации  $\phi$ , бисериальный коэффициент корреляции  $r_{pb}$ , рангово-бисериальный коэффициент корреляции  $r_{rb}$ , корреляционное отношение Пирсона  $\eta$ . Последний коэффициент корреляции позволяет обнаруживать нелинейную связь между переменными, измеренными в шкале интервалов. Остальные служат для обнаружения линейной связи между переменными, измеренными в различных шкалах (табл. 7.19).

**Таблица 7.19 Меры связи для различных типов шкал измерения переменных**

Тип шкалы		Мера связи
Переменная X	Переменная Y	
Интервальная	Интервальная	Коэффициент корреляции Пирсона $r_{xy}$ , коэффициент ранговой корреляции Спирмена $r_s$ , корреляционное отношение Пирсона $\eta$
Ранговая	Ранговая	Коэффициент ранговой корреляции Спирмена $r_s$ , $\tau$ Кендалла
Дихотомическая	Дихотомическая	Коэффициент ассоциации $\phi$
Дихотомическая	Ранговая	Рангово-бисериальный $r_{rb}$
Дихотомическая	Интервальная	Бисериальный $r_{pb}$

Величина коэффициента корреляции всегда больше  $-1$  и меньше  $+1$ . Если значение коэффициента корреляции близко к  $+1$ , это озна-

часть, что более высоким значениям переменной  $X$  соответствуют более высокие значения переменной  $Y$ . Если же значение коэффициента корреляции близко к  $-1$ , это означает, что более высоким значениям переменной  $X$  соответствуют более низкие значения переменной  $Y$ . В тех случаях, когда соответствие между величиной значений переменных  $X$  и  $Y$  отсутствует, значение коэффициента корреляции близко к нулю.

Традиционно сложились две системы классификации корреляционных связей: общая и частная. Существование общей классификации корреляционной связи определено тем, что в отличие от всех остальных статистических критериев абсолютное значение коэффициента корреляции не превосходит единицы. При общей классификации (табл. 7.20) тип корреляционной связи определяется по величине коэффициента корреляции,

**Таблица 7.20. Общая классификация корреляционных связей**

Тип корреляционной связи	Направленность корреляционной связи	Значение коэффициента корреляции
Сильная	Положительная	$0,7 < r < 1,0$
Средняя		$0,5 < r < 0,69$
Умеренная		$0,3 < r < 0,49$
Слабая		$0,2 < r < 0,29$
Очень слабая		$0 < r < 0,19$
Сильная	Отрицательная	$-1,0 < r < -0,7$
Средняя		$-0,69 < r < -0,5$
Умеренная		$-0,49 < r < -0,3$
Слабая		$-0,29 < r < -0,2$
Очень слабая		$-0,19 < r < 0$

При частной классификации корреляционных связей тип корреляционной связи определяется в зависимости уровня значимости достигает величина коэффициента корреляции при данном объёме выборки (табл. 7.21).

Чем больше объём выборки, тем меньшей величины коэффициента корреляции оказывается достаточно, чтобы корреляция была признана статистически значимой. В результате при малом объёме выборки может оказаться так, что сильная корреляция может ока-

заться незначимой. В то же время при больших объемах выборки даже слабая корреляция может оказаться достоверной.

**Таблица 7.21. Частная классификация корреляционных связей**

Тип корреляционной связи	Значение коэффициента корреляции	Уровень статистической значимости
Высокая значимая корреляция	$r_{\text{эмп}} \geq r_{\text{кр}}^{0,01}$	0,01
Значимая корреляция	$r_{\text{кр}}^{0,05} \leq r_{\text{эмп}} < r_{\text{кр}}^{0,01}$	0,05
Тенденция достоверной связи	$r_{\text{кр}}^{0,1} \leq r_{\text{эмп}} < r_{\text{кр}}^{0,05}$	0,1
Незначимая корреляция	$r_{\text{эмп}} < r_{\text{кр}}^{0,1}$	–

Нами будут рассмотрены правила вычисления наиболее часто используемых коэффициентов корреляции К.Пирсона и Ч.Спирмена. Остальные меры связи подробно описаны в руководстве [22].

### **Коэффициент корреляции $r_{xy}$ Пирсона**

Коэффициент корреляции Пирсона служит для выявления степени согласованности изменений двух переменных  $X$  и  $Y$ , измеренных в интервальной шкале.

#### **Условия применения коэффициента корреляции Пирсона**

- обе переменные должны быть измерены по шкале интервалов (т.е. в шкалах стандартных показателей);
- данные должны быть распределены по нормальному закону.

Эмпирическое значение коэффициента корреляции Пирсона вычисляется по формуле (7.45).

$$r_{xy} = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum (X_i - \bar{X})^2 \sum (Y_i - \bar{Y})^2}}, \quad (7.45)$$



где  $\bar{X}$  и  $\bar{Y}$  – средние арифметические значения переменных  $X$  и  $Y$  соответственно.

После того, как вычислено эмпирическое значение коэффициента корреляции, статистическое решение о степени согласованности изменений переменных  $X$  и  $Y$  может приниматься по одной из двух систем классификации корреляционных связей. Выбор системы классификации производится в соответствии с задачами экспериментального исследования.

### **Коэффициент ранговой корреляции $r_s$ Спирмена**

Если при вычислении коэффициента корреляции Пирсона используются фактические значения признаков, то при вычислении коэффициента ранговой корреляции Спирмена – их ранги. Таким образом, для подсчёта коэффициента ранговой корреляции необходимо располагать двумя рядами значений, которые могут быть проранжированы, поэтому коэффициент ранговой корреляции более универсален, чем коэффициент корреляции Пирсона, т.к. его можно применять к любым численно измеренным данным.

#### **Условия применения коэффициента корреляции Спирмена**

- обе переменные должны быть измерены в одинаковой шкале (порядковой или интервальной);
- по каждой переменной должно быть не менее 5 значений.

#### **Правила расчёта коэффициента корреляции Спирмена**

- Проранжировать испытуемых в соответствии со значением переменной  $X$ , приписывая меньшему значению меньший ранг. Занести в таблицу полученные значения рангов  $R_X$ , записав их отдельным столбцом.
- Проранжировать испытуемых (при ранжировании особое внимание следует обращать на сохранение всех горизонтальных связей в таблице данных) в соответствии со значением переменной  $Y$ , приписывая меньшему значению меньший ранг. Занести в таблицу полученные значения рангов  $R_Y$ , записав их отдельным столбцом.

- Подсчитать разности  $d$  между рангами  $R_X$  и  $R_Y$ . Внести полученные результаты в соответствующий столбец таблицы.
- Возвести каждую разность в квадрат и заполнить следующий столбец таблицы. Подсчитать сумму квадратов.
- При наличии одинаковых рангов рассчитать поправки:

$$T_x = \sum (a^3 - a)/12; T_y = \sum (e^3 - e)/12, \quad (7.46)$$

где  $a$  – объём каждой группы одинаковых рангов в ряду  $X$ ;

$e$  – объём каждой группы одинаковых рангов в ряду  $Y$ .

- При отсутствии одинаковых рангов вычислить коэффициент ранговой корреляции по формуле (7.47):

$$r_s = 1 - \frac{6 \sum d^2}{n(n^2 - 1)}. \quad (7.47)$$

- При наличии одинаковых рангов вычислить коэффициент ранговой корреляции по формуле

$$r_s = 1 - \frac{6 \cdot \sum d^2 + T_x + T_y}{n(n^2 - 1)}. \quad (7.48)$$

- Определить критические значения для данного числа испытуемых по таблице. Сделать вывод о принятии гипотезы.

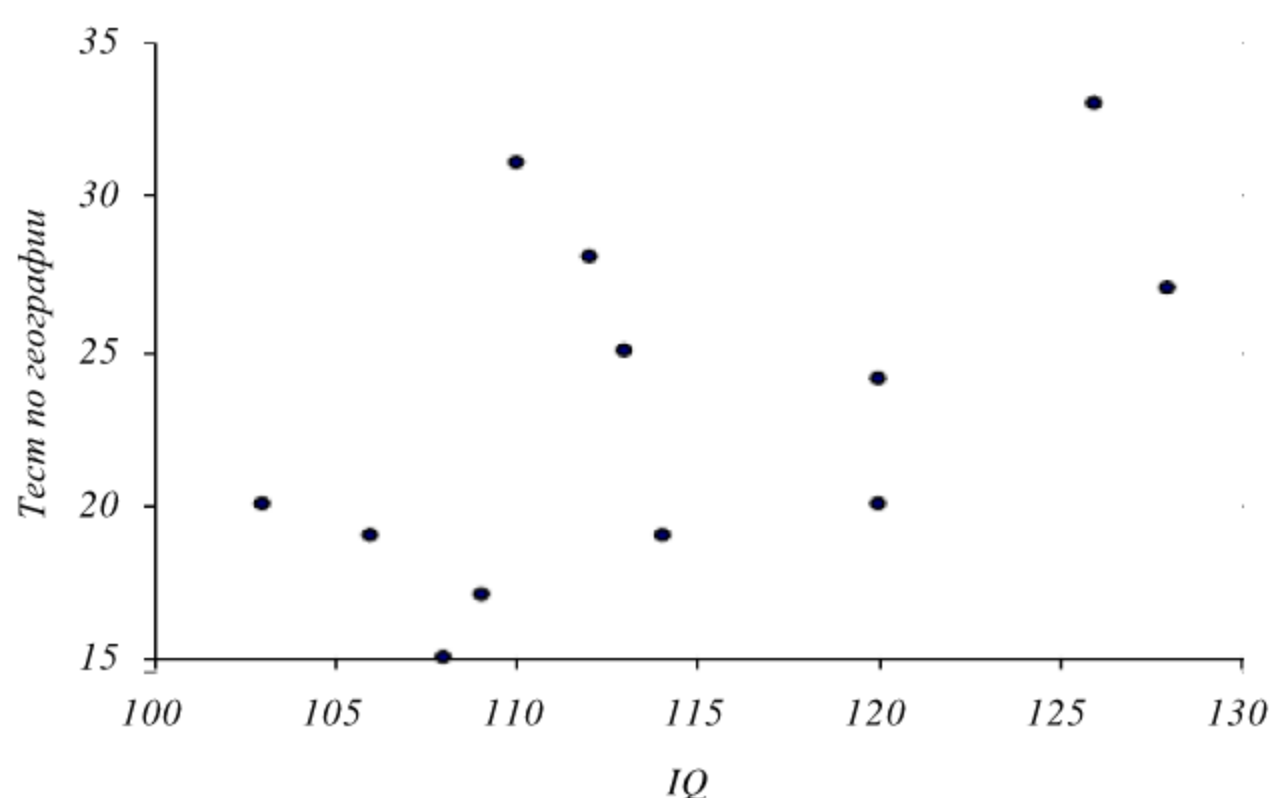
**Пример.** Каким образом связана успеваемость по различным школьным предметам с коэффициентом интеллекта школьников? Для ответа на этот вопрос двенадцати шестиклассникам были предложены тесты школьных достижений по географии и математике, состоящие из 35 вопросов. Эти же дети были протестированы с помощью теста Векслера. Результаты обследования представлены в таблице 7.22.

Сначала изобразим связь между IQ и оценками по географии и математике с помощью точечных диаграмм, или диаграмм рассеяния. Число точек на каждой диаграмме соответствует объёму выборки (в нашем случае  $n=12$ ), координаты каждой точки равны значениям IQ и результату тестирования по предмету у соответствующего испытуемого (рис. 7.12, 7.13). На рис. 7.12 точки распо-

ложены хаотично, что приводит нас к предположению об отсутствии значимой корреляционной связи между IQ ученика и его успеваемостью по географии.

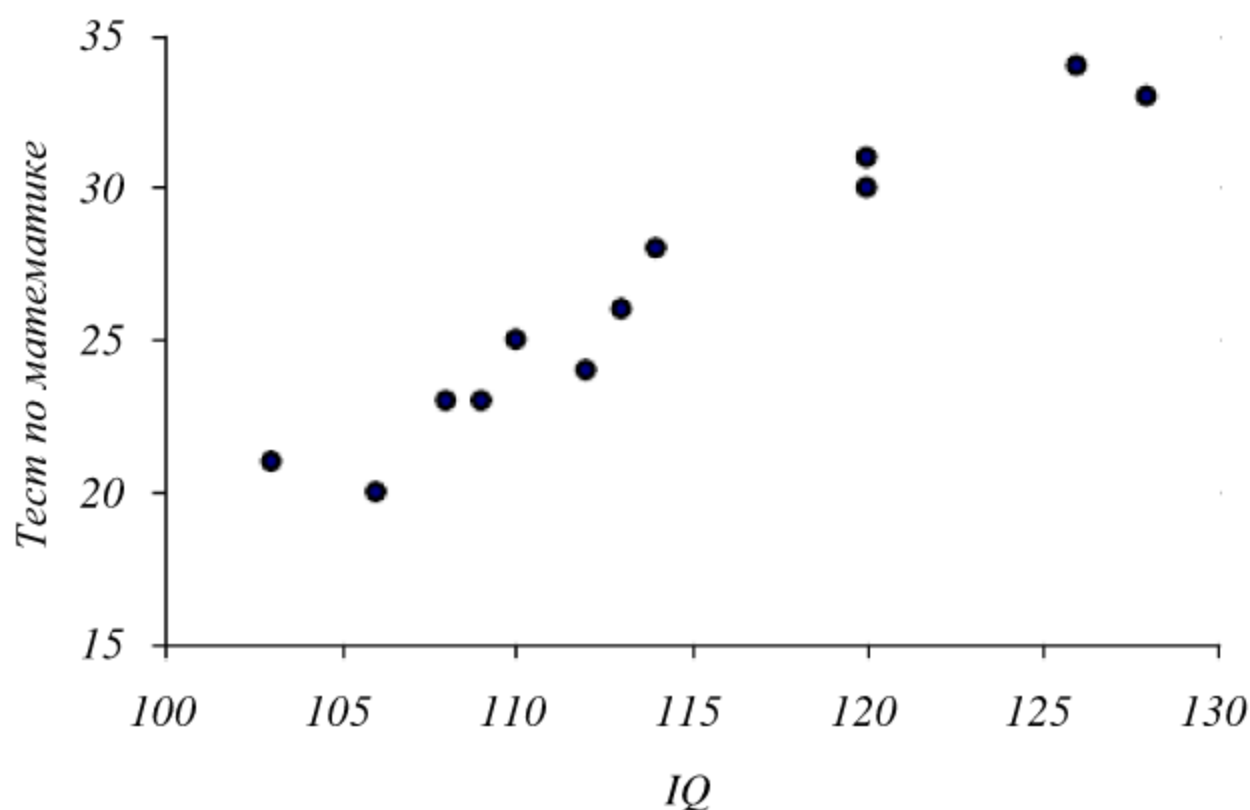
**Таблица 7.22. Значение IQ и результаты тестов школьных достижений у испытуемых**

№	IQ	Тест по географии	Тест по математике
1	120	31	31
2	112	25	24
3	110	19	25
4	120	24	30
5	103	17	21
6	126	28	34
7	113	18	26
8	114	20	28
9	106	16	20
10	108	15	23
11	128	27	33
12	109	19	23



**Рис.7.12 Графическое изображение связи IQ с успеваемостью по географии**

На рис. 7.13 точки сгруппированы вдоль некоторой прямой, что позволяет нам выдвинуть гипотезу о наличии достоверной корреляционной связи между IQ ученика и его успеваемостью по математике.



**Рис.7.13. Графическое изображение связи IQ с успеваемостью по математике**

#### **Статистические гипотезы:**

$H_0$ — между IQ учеников и успеваемостью по математике не существует значимой корреляционной связи;

$H_1$ — между IQ учеников и успеваемостью по математике существует значимая положительная корреляционная связь.

Проверим выдвинутые статистические гипотезы с помощью коэффициента ранговой корреляции Спирмена.

Сначала занесём в таблицу индивидуальные результаты по тесту Векслера и тесту по математике и упорядочим их по возрастанию IQ испытуемых (табл. 7.23). Проранжируем значения IQ, приписывая меньшему значению меньший ранг.

Далее упорядочим все данные по возрастанию результата теста по математике и также проранжируем эти результаты (табл. 7.24). После этого вычислим разности  $d$  между рангами по IQ и математике у каждого испытуемого, возведём эти разности в квадрат и просуммируем (вычислим  $\sum d^2$ ).

**Таблица 7.23. Результат ранжирования испытуемых по величине IQ**

№ испытуемого	IQ	Ранг IQ	Тест по математике
11	128	12	33
6	126	11	34
1	120	9,5	31
4	120	9,5	30
8	114	8	28
7	113	7	26
2	112	6	24
3	110	5	25
12	109	4	23
10	108	3	23
9	106	2	20
5	103	1	21

Рассчитаем поправки на одинаковые ранги. В обоих ранговых рядах присутствует всего по одной группе одинаковых рангов:  $a = 2$ ;  $b = 2$ . Таким образом,

$$T_{IQ} = (a^3 - a)/12 = 0,5; T_m = (b^3 - b)/12 = 0,5. \quad (7.48)$$

**Таблица 7.24. Ранжирование испытуемых по результату теста по математике**

№ испытуемого	IQ	Ранг IQ	Тест по матем.	Ранг по матем.	d	d <sup>2</sup>
6	126	11	34	12	-1	1
11	128	12	33	11	1	1
1	120	9,5	31	10	-0,5	0,25
4	120	9,5	30	9	0,5	0,25
8	114	8	28	8	0	0
7	113	7	26	7	0	0
3	110	5	25	6	-1	1
2	112	6	24	5	1	1
12	109	4	23	3,5	0,5	0,25
10	108	3	23	3,5	-0,5	0,25
5	103	1	21	2	-1	1
9	106	2	20	1	1	1
					$\Sigma d^2$	<b>7</b>

Значение коэффициента корреляции Спирмена

$$r_s = 1 - \frac{6 \cdot \sum d^2 + T_{IQ} + T_M}{n(n^2 - 1)} = 1 - \frac{6 \cdot 7 + 0,5 + 0,5}{12 \cdot (144 - 1)} = 0,975 \quad (7.49)$$

Критические значения коэффициента корреляции Спирмена для объёма выборки  $n=12$

$$r_s^{кр} = \begin{cases} 0,58(p \leq 0.05) \\ 0,73(p \leq 0.01) \end{cases} \quad (7.50)$$

Сопоставив эмпирическое значение коэффициента корреляции с критическими, делаем вывод о высокой значимой корреляционной связи между IQ шестиклассников с их успеваемостью по математике. Отметим, что в соответствии с общей системой классификации корреляционных связей получается тот же вывод.

Аналогичным образом вычисляется значение коэффициента корреляции Спирмена между величиной IQ и результатом теста по географии.

### **Интерпретация коэффициентов корреляции.**

Наличие статистически значимой корреляционной связи между переменными можно интерпретировать одним из следующих способов.

- Уровень значений одной переменной непосредственно соответствует уровню другой (например, уровень мотивации и эффективность выполнения задачи).
- Связь между переменными обусловлена третьей переменной, которая в ходе эксперимента не измерялась (например, связь между уровнем развития различных познавательных процессов, в соответствии с теорией Ч.Спирмена обусловлена величиной G-фактора).
- Корреляция, обусловленная неоднородностью выборки (например, в выборке, состоящей из мужчин-программистов и женщин-журналистов связь между полом и экстраверсией-интроверсией на самом деле может быть профессионально обусловленной). Для грамотной интерпретации коэффициентов



корреляции данных, полученных при обследовании неоднородных выборок испытуемых, необходимо провести сравнительный анализ по каждой переменной с помощью статистических критериев различий либо с помощью дисперсионного анализа, а затем уже принимать решение о вычислении коэффициента корреляции по объединённой выборке либо по каждой группе отдельно.

- Если размах вариативности одной из переменных в обследованной выборке слишком узкий, то нельзя распространять полученную корреляцию на весь возможный диапазон значений. В подобных случаях возможны самостоятельные выводы по различным диапазонам исследуемого признака. Так, сила связи между IQ и креативностью различна для различных показателей IQ. В исследованиях П.Торренса показано, что при  $IQ < 120$  существует достоверная корреляционная связь между IQ и креативностью, а при  $IQ > 130$  такой связи нет.

В заключение заметим, что причины взаимосвязи переменных в психологии, педагогике и общественных науках настолько сложны, что интерпретация в большинстве случаев неочевидна и неоднозначна.

### **Контрольные вопросы и задания**

1. *Каким образом можно графически изобразить связь между двумя переменными? Сделайте иллюстрацию к случаю, когда между переменными имеется слабая обратная связь.*
2. *Охарактеризуйте суть общей и частной систем классификации корреляционных связей. В чём различие между ними?*
3. *Какой из методов вычисления коэффициента корреляции (Пирсона или Спирмена) более универсален и почему?*
4. *Между переменными есть сильная связь, но коэффициент корреляции Пирсона оказался близким к нулю. Может ли такое быть и если да, то в каких случаях?*
5. *Какой из коэффициентов корреляции (Пирсона или Спирмена) в большей степени подходит для установления причинно-следственной зависимости между переменными, т.е. позволяет достоверно опреде-*

лить, какая из переменных является причиной, а какая следствием при наличии достоверной связи между ними?

6. Вычислите значение коэффициента корреляции Спирмена между величиной IQ и результатом теста по географии (табл. 7.20). Проинтерпретируйте полученные значения коэффициентов корреляции. Почему между величиной IQ и результатом теста по математике существует более сильная связь?

7. В эксперименте по изучению невербального и вербального интеллекта приняли участие 20 школьников (табл. 7.25). Измерялось время решения отдельных субтестов теста Векслера в секундах. Существует ли статистически достоверная взаимосвязь между временем выполнения вербальных и невербальных субтестов? Применить два критерия для доказательства гипотезы (параметрический, проверив правомочность его применения, и непараметрический).

**Таблица 7.25. Исходные данные**

№	Невербальные субтесты	Вербальные субтесты
1	19	17
2	32	7
3	33	17
4	44	28
5	28	27
6	35	31
7	39	20
8	39	17
9	44	35
10	44	43
11	24	10
12	37	28
13	29	13
14	40	43
15	32	45
16	42	24
17	48	45
18	42	26
19	33	16
20	47	26

**Требования к оформлению задач см. в заданиях к параграфу 7.2.**

## 7.5. Многофункциональные статистические критерии

Многофункциональные статистические критерии имеют широкий спектр применения и минимум ограничений. Так, данные могут быть представлены в любой шкале, выборки могут быть как независимыми (сравниваются две различные выборки испытуемых), так и связными (сопоставляются показатели одной и той же выборки в различных условиях). Но, с другой стороны, вследствие того, что по сравнению с другими критериями применение многофункциональных критериев гораздо менее формализовано, перед их использованием требуется более основательный предварительный анализ данных (либо визуальный, либо с применением других критериев, например, критерия Колмогорова-Смирнова).

### **Критерий $\varphi^*$ – угловое преобразование Фишера.**

Критерий  $\varphi^*$  предназначен для оценки различий между процентными долями двух выборок по частоте встречаемости определённых значений признака, которые могут определяться как количественно (результат выполнения теста, превышающий определённый балл), так и качественно (выбор определённого пункта анкеты, категория респондентов и т.п.).

Обозначим объёмы сопоставляемых выборок  $n_1$  и  $n_2$ , а число испытуемых в каждой выборке, показавших интересующий нас результат, соответственно  $n_1^*$  и  $n_2^*$ .

### **Условия применения критерия $\varphi^*$ :**

- $n_1^* \neq 0$  и  $n_2^* \neq 0$ , т.е. в каждой из групп должен быть хотя бы один испытуемый, показавший интересующий нас результат;
- При  $n_1 \leq 5$  и  $n_2 \leq 5$  существуют ограничения на объёмы выборок, представленные в табл. 7.26.

**Таблица 7.26 Нижние границы объёмов сопоставляемых выборок**

$n_1$	$n_2$
2	$\geq 30$
3	$\geq 7$
4	$\geq 5$
5	$\geq 5$

**Статистические гипотезы:**

$H_0$  – частота встречаемости признака в выборке 1 не превышает частоты встречаемости признака в выборке 2;

$H_1$  – частота встречаемости признака в выборке 1 превышает частоту встречаемости признака в выборке 2.

**Правила расчёта критерия  $\varphi^*$ .**

- Вычислить  $p_1$  и  $p_2$  – процентные доли испытуемых в сопоставляемых выборках, показавших интересующий нас результат:

$$p_1 = \frac{n_1^*}{n_1}; p_2 = \frac{n_2^*}{n_2}. \quad (7.51)$$

- Вычислить величины  $\varphi_1$  и  $\varphi_2$ , соответствующие сопоставляемым процентным долям  $p_1$  и  $p_2$ , в которых присутствует интересующий нас эффект:

$$\varphi_1 = 2 \arcsin \sqrt{p_1}; \varphi_2 = 2 \arcsin \sqrt{p_2}. \quad (7.52)$$

Вычисление функции  $\arcsin$  можно производить как в электронных таблицах Excel, так и с помощью специальных таблиц (см. Приложение).

- Эмпирическое значение критерия вычисляется по формуле

$$\varphi_{\text{эмп}}^* = (\varphi_1 - \varphi_2) \cdot \sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}}. \quad (7.53)$$

- Полученное  $\varphi_{\text{эмп}}^*$  сопоставляется с критическими значениями

$$\varphi_{\text{кр}}^* = \begin{cases} 1,64(p \leq 0.05) \\ 2,31(p \leq 0.01) \end{cases}. \quad (7.54)$$

При  $\varphi_{\text{эмп}}^* > \varphi_{\text{кр}}^*$   $H_0$  отвергается.

При необходимости можно определить точный уровень статистической значимости полученного  $\varphi_{\text{эмп}}^*$  по таблице (см. Приложение 2).

Рассмотрим несколько примеров вычисления углового преобразования Фишера.

**Пример 1.** В главе 6 нами были рассмотрены результаты диагностики когнитивной сферы группы учащихся Обнинского политехникума.

На основе результатов диагностики и оценки уровня развития когнитивной сферы была построена таблица сопряжённости (табл. 7.27). Можно ли утверждать, что доля мужчин с высоким уровнем развития когнитивной сферы достоверно выше, чем доля женщин?

**Таблица 7.27. Таблица сопряженности**

Уровень развития когнитивной сферы	Пол	
	М	Ж
В	6 (60%)	2 (40%)
Н	4 (40%)	3 (60%)

**Статистические гипотезы:**

$H_0$  – доля испытуемых с высоким уровнем развития когнитивной сферы в выборке мужчин не выше, чем в выборке женщин;

$H_1$  – доля испытуемых с высоким уровнем развития когнитивной сферы в выборке мужчин выше, чем в выборке женщин.

В соответствии с правилами применения углового преобразования Фишера обозначим объёмы сопоставляемых выборок  $n_1 = 10$  (мужчины) и  $n_2 = 5$  (женщины), а число испытуемых в каждой из выборок, показавших интересующий нас результат, соответственно

$$n_1^* = 6 \text{ и } n_2^* = 2. \quad (7.55)$$

Вычислим процентные доли испытуемых в сопоставляемых выборках, показавших высокий результат:

$$p_1 = \frac{6}{10} = 0,6; \quad p_2 = \frac{2}{5} = 0,4. \quad (7.56)$$

Вычислим величины  $\varphi_1$  и  $\varphi_2$ :

$$\varphi_1 = 2 \arcsin \sqrt{p_1} = 1,77 \quad (7.57)$$

$$\varphi_2 = 2 \arcsin \sqrt{p_2} = 1,37$$

Эмпирическое значение критерия равно

$$\varphi_{\text{эмп}}^* = (\varphi_1 - \varphi_2) \cdot \sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}} = 0,4 \cdot \sqrt{\frac{10 \cdot 5}{10 + 5}} = 0,73 \quad (7.58)$$

Сопоставим полученное значение  $\varphi_{\text{эмп}}^*$  с критическими значениями

$$\varphi_{\text{кр}}^* = \begin{cases} 1,64 (p \leq 0.05) \\ 2,31 (p \leq 0.01) \end{cases} \quad (7.59)$$

Так как  $\varphi_{\text{эмп}}^* < \varphi_{\text{кр}}^*$ , принимается гипотеза  $H_0$ : статистически значимых гендерных различий в категории испытуемых с высоким уровнем развития когнитивной сферы не существует.

Отметим, что в рассмотренном примере угловое преобразование Фишера применялось для сопоставления данных, представленных в номинативной шкале.

В следующем примере данные представлены в шкале порядка.

**Пример 2.** Существует ли статистически значимая связь роста детей с ростом их родителей на основе данных Ф.Гальтона [22, с. 38]? Данные представлены в табл. 7.28.

Сначала вычислим относительные частоты встречаемости детей различного роста у каждой категории родителей (табл. 7.29) и построим гистограмму распределения частот (рис. 7.14).

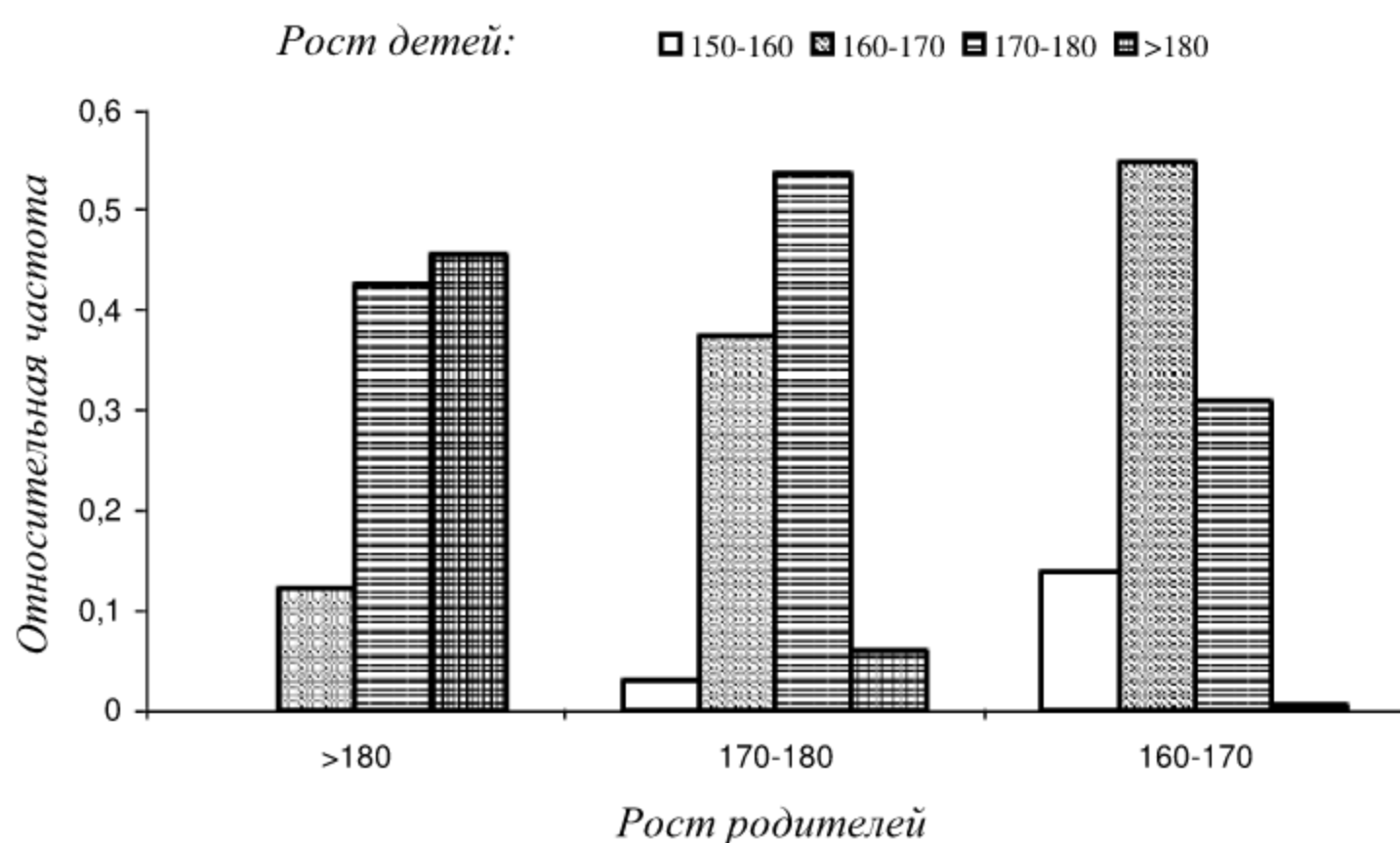


**Таблица 7.28. Соотношение роста детей и родителей**

Показатели		Рост детей				всего
		150-160	160-170	170-180	181 и выше	
Рост родителей	180 и выше	0	8	28	30	66
	170-180	19	255	366	41	681
	160-170	25	99	56	1	181
всего		44	362	450	72	928

**Таблица 7.29. Таблица относительных частот**

Показатели		Рост детей			
		150-160	160-170	170-180	181 и выше
Рост родителей	180 и выше	0	0,12	0,42	0,45
	170-180	0,03	0,37	0,54	0,06
	160-170	0,14	0,55	0,31	0,01



**Рис. 7.14. Распределение роста детей у различных категорий родителей**

Из табл. 7.29 видно, что при росте родителей 181 см и выше в 45% случаев рост детей также выше 180 см, в 42% случаев – от 170 до 180 см и лишь в 12% случаев – ниже 170 см. Низкорослых детей (менее 160 см) у родителей высокого роста в данной выборке нет. Если же рост родителей находится в диапазоне 160–170 см, то в 55% случаев рост их детей находится в том же диапазоне, в 31% случаев – от 170 до 180 см, в 14% случаев – от 150 до 160 см и лишь в 1% случаев превышает 181 см.

Проверим статистическую значимость различий частоты встречаемости детей ростом 160–170 см в выборках родителей ростом 170–180 см и выше 180 см с помощью углового преобразования Фишера.

**Статистические гипотезы:**

$H_0$  – доля детей ростом 160–170 см в выборке родителей ростом 170–180 см не выше, чем доля детей этого роста в выборке родителей, рост которых более 180 см;

$H_1$  – доля детей ростом 160–170 см в выборке родителей ростом 170–180 см достоверно выше, чем доля детей этого роста в выборке родителей, рост которых более 180 см.

В соответствии с правилами применения углового преобразования Фишера обозначим объёмы сопоставляемых выборок (см. табл. 7.28)  $n_1 = 681$  (число семей, в которых рост родителей 170–180 см) и  $n_2 = 66$  (число семей, в которых рост родителей более 180 см), а число детей ростом 160–170 см в каждой из выборок соответственно  $n_1^* = 255$  и  $n_2^* = 8$ .

Сопоставляемые процентные доли соответственно равны (табл. 7.29)  $p_1 = 0,37$ ;  $p_2 = 0,12$ .

Вычислим величины  $\varphi_1$  и  $\varphi_2$ :

$$\varphi_1 = 2 \arcsin \sqrt{p_1} = 1,307; \quad \varphi_2 = 2 \arcsin \sqrt{p_2} = 0,707. \quad (7.60)$$

Эмпирическое значение критерия равно

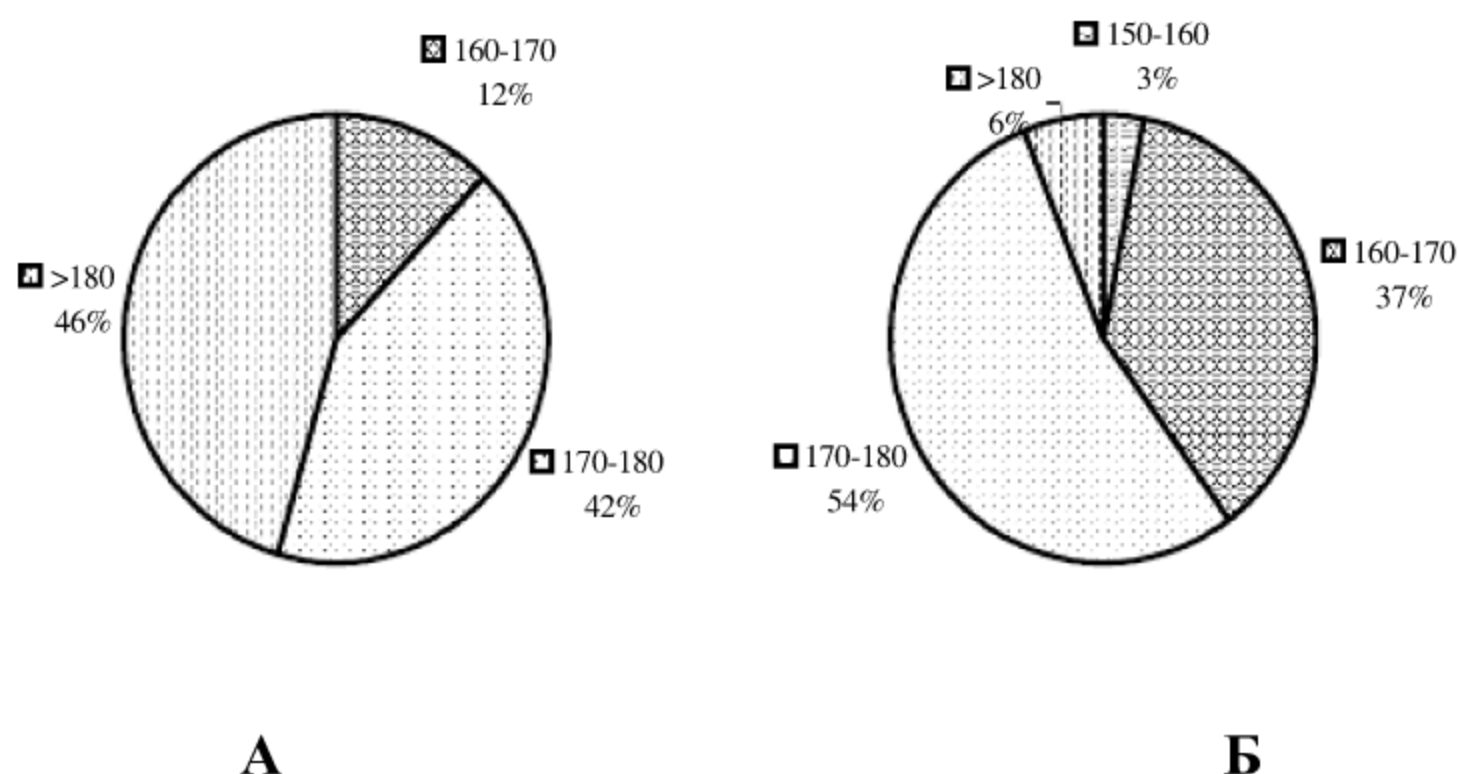
$$\varphi_{\text{эмп}}^* = (\varphi_1 - \varphi_2) \cdot \sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}} = (1,307 - 0,707) \cdot \sqrt{\frac{681 \cdot 66}{681 + 66}} = 4,66 \quad (7.61)$$

Сопоставим полученное значение  $\varphi_{\text{эмп}}^*$  с критическими значениями

$$\varphi_{\text{кр}}^* = \begin{cases} 1,64 (p \leq 0.05) \\ 2,31 (p \leq 0.01) \end{cases} \quad (7.62)$$

Поскольку  $\varphi_{\text{эмп}}^* > \varphi_{\text{кр}}^*$ , принимается гипотеза  $H_1$  ( $p < 0,01$ ).

Доказанные различия можно проиллюстрировать с помощью круговой диаграммы (рис. 7.15).



**Рис.7.15. Процентные доли детей различного роста у родителей выше 180 см (А) и у родителей 170-180 см (Б)**

В двух рассмотренных примерах точки разделения шкалы значений на категории выбирались исследователем «вручную» на основе визуального анализа исходных данных или исходя из каких-либо других соображений. Заметим, что правильный выбор точки разделения на категории повышает мощность критерия  $\varphi^*$ . Для наиболее точного определения местоположения границы разделения шкалы значений на категории используется критерий Колмогорова-Смирнова.

### Контрольные вопросы и задания

1. В примере 1 (табл. 7.27) увеличьте объёмы выборки в 10 раз
2. На основе данных Ф.Гальтона (табл. 7.28, 7.29) проверьте статистическую значимость различий долей детей других ростовых категорий у родителей ростом 160–170, 170–180 и более 180 см.

**Таблица 7.30. Исходные данные**

		Пол	
		М	Ж
Уровень IQ	В	60 (60%)	20 (40%)
	Н	40 (40%)	30 (60%)

*Требования к оформлению представлены в конце параграфа 7.2.*

## **Глава 8.**

### **Обзор методов многомерного анализа данных**

В настоящей главе представлен обзор аналитико-статистических методов обработки многомерных психодиагностических данных. Подробное изложение этих методов описано в специализированной математической литературе. Целью обзорного изложения основных многомерных методов является информирование специалиста-психолога о возможностях их использования для анализа больших массивов данных.

При использовании методов многомерного анализа особенно важно помнить о следующих правилах получения исходных психодиагностических данных:

- экспериментальное обследование испытуемых должно быть тщательно спланировано и проведено с соблюдением всех норм регламентации;
- измерение должно проводиться с помощью надёжной и валидной психодиагностической методики (либо детально продуманной системы экспертных оценок), выбор которой произведён на основе глубокой теоретической проработки изучаемого психологического явления;
- объём выборки должен быть не менее 100 человек;
- сырые баллы должны переводиться в стандартные показатели (Т-баллы, стеноны) только в том случае, когда обследованная выборка испытуемых принадлежит к той же генеральной совокупности, что и выборка стандартизации;
- перед применением аналитико-статистических методов исходные данные должны быть нормированы (переведены в Z-оценки).

В главе будут рассмотрены следующие типы задач, решаемых с помощью аналитико-статистических методов:

- Изучить зависимость изменений значений зависимой переменной  $Y$  от одной или нескольких независимых переменных  $X$  или предсказать, какими будут значения  $Y$  при определённых значениях  $X$ . В данном случае применяют **регрессионный анализ**.
- Изучить влияние отдельных переменных, а также их сочетаний на изменчивость изучаемого признака. Решается с помощью **дисперсионного анализа**.
- Провести классификацию испытуемых по результатам измерения нескольких показателей. Для этого используются многомерные статистические методы: **дискриминантный анализ** и **кластерный анализ**.
- Определить структуру взаимосвязей между переменными, сократить их число и перейти от совокупности непосредственно измеряемых переменных к комплексным обобщённым факторам, за которыми стоят комбинации этих исходных переменных. Такого рода проблемы решаются с помощью методов **факторного анализа**.

## 8.1. Регрессионный анализ

При анализе данных психологических исследований существует определённый тип задач, называемых задачами предсказания и оценивания.

- Как и насколько хорошо можно **предсказать** успеваемость по общей психологии в вузе по результатам ЕГЭ абитуриентов?
- Насколько правильно можно **предсказать** успешность профессиональной деятельности выпускников вуза по результатам оценки профессионально важных личностных характеристик на третьем курсе?
- Как точно можно **предсказать** заработок в 35 лет по рейтинговым оценкам в вузе?

Чтобы вывести способ оценивания объекта по одной переменной (мы обозначим её  $Y$ ) на основе другой переменной ( $X$ ), необ-



ходимо знать, как связаны между собой  $X$  и  $Y$ . Переменная  $Y$ , которую требуется оценить, называется **зависимой переменной** или **откликом**, а переменная  $X$ , используемая для её оценки – **независимой переменной** или **предиктором**, или **фактором**. Эта терминология является отражением того факта, что в данном случае ставится задача определения **зависимости**  $Y$  от  $X$ , или **предсказания**, какими будут значения  $Y$  при данных значениях  $X$ . Такого рода задачи решаются с помощью регрессионного анализа.

Таким образом, суть регрессионного анализа заключается в нахождении функции

$$Y = F(X), \quad (8.1)$$

которая называется **функцией предсказания**.

**Рассмотрим пример.** Требуется предсказать результат выполнения студентами-психологами 3 курса контрольной работы по математическим методам в психологии (зависимая переменная  $Y$ ) по результатам интеллектуального теста Айзенка в начале учебного года (независимая переменная  $X$ ).

Для решения этой задачи необходимо предварительно собрать данные для выборки студентов-психологов 3 курса:

- а) провести диагностику интеллектуальных способностей с помощью теса Айзенка в начале учебного года;
- б) провести контрольную работу по математическим методам в психологии в конце учебного года. Хорошо, если указанные данные будут собраны за несколько лет.

Допустим, что известны результаты 20 студентов, на основе которых нужно вывести интересующую нас функцию предсказания. Сразу оговоримся, что объём выборки ограничен количеством 20 человек исключительно в целях наглядности и простоты восприятия данного примера. В реальных исследованиях для построения функции предсказания выборка должна быть не менее 100 чел.

Контрольная работа представляет собой тест достижений, включающий 50 вопросов. Результаты теста Айзенка переведены в стандартную IQ шкалу (табл. 8.1).

**Таблица 8.1. Результаты обследования**

№	X	Y		№	X	Y
1	95	33		11	110	41
2	100	31		12	110	44
3	100	35		13	111	40
4	102	38		14	112	45
5	103	41		15	112	48
6	105	37		16	114	45
7	106	37		17	114	49
8	106	39		18	115	47
9	106	43		19	117	43
10	109	40		20	118	48

На основе данных таблицы 8.1 выведем функцию предсказания, связывающее X и Y в этой группе. Впоследствии это уравнение можно использовать для студентов, значения X которых нам известны, а Y нам хотелось бы оценить.

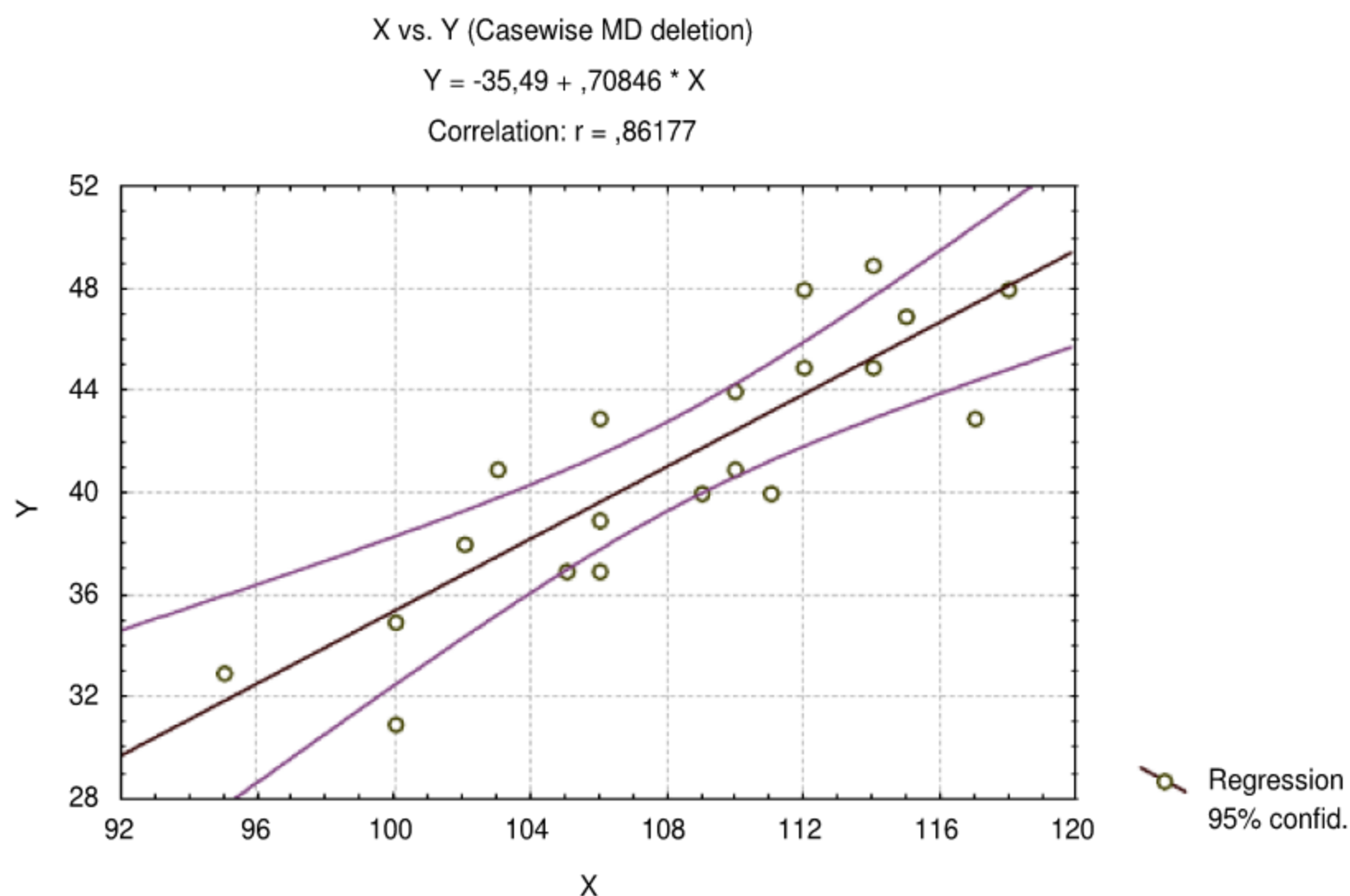
На первом этапе вычислим коэффициент корреляции Пирсона между переменными X и Y, который в рассматриваемом случае принимает значение  $r=0,86$ . Величина коэффициента корреляции свидетельствует о сильной прямой связи переменных на достоверном уровне статистической значимости ( $p<0.001$ ).

На рис. 8.1 представлена диаграмма рассеяния, на которой видно, что все точки выстраиваются вдоль прямой линии, т.е. функция предсказания является линейной. Она называется **линией предсказания или линией регрессии**.

Как известно, в общем виде уравнение линейной функции имеет вид:

$$Y = b_1 X + b_0 . \quad (8.2)$$

Для того, чтобы определить значения  $b_1$  и  $b_0$ , используется метод наименьших квадратов.



**Рис. 8.1. Диаграмма рассеяния**

Каждому значению  $X_i$  соответствует точка на линии регрессии, значение ординаты которой

$$\tilde{Y}_i = b_1 X_i + b_0. \quad (8.3)$$

В тоже время, исходя из таблицы исходных значений, каждому значению  $X_i$  соответствует вполне определённое значение  $Y_i$ .

$$\tilde{Y}_i = Y_i + e_i, \quad (8.4)$$

где  $e_i$  – ошибка оценивания для  $i$ -го испытуемого.

Таким образом,

$$e_i = Y_i - \tilde{Y}_i \quad (8.5)$$

– это ошибка оценки.

При вычислении коэффициентов линии регрессии  $b_0$  и  $b_1$  определяются таким образом, чтобы сумма квадратов ошибок оценки  $\sum e_i^2$  была минимальной. Критерий наименьших квадратов был изобретён К.Гауссом и в той или иной форме лежит в основе многих теоретических и прикладных статистических работ. Не оста-

навливаясь на процедуре вычисления  $b_0$  и  $b_1$  с помощью метода наименьших квадратов, отметим, что  $b_0$  и  $b_1$  связаны со значениями средних значений и стандартных отклонений обоих распределений, а также с коэффициентом корреляции Пирсона между переменными  $X$  и  $Y$  и вычисляются по формулам (8.5), (8.6).

$$b_1 = r_{xy} \frac{\sigma_y}{\sigma_x}; \quad (8.5)$$

$$b_0 = \bar{Y} - b_1 \bar{X}. \quad (8.6)$$

При построении диаграммы рассеяния в программе Statistica (рис.8.1), на графике по умолчанию изображается линия регрессии, а в строке под заголовком – уравнение линии регрессии.

Заметим, что теоретически линию регрессии можно построить для любого множества значений признаков. Понятно, что если между переменными отсутствует линейная связь, то предсказание значений  $Y$  по значениям  $X$  с помощью такой линии регрессии будет недостоверным. Для проверки уровня значимости коэффициентов регрессии вычисляется  $t$ -критерий Стьюдента.

$$t = \frac{b_1}{\sqrt{\frac{(1-r_{xy}^2) \cdot \sum (x_i - \bar{x})^2}{(n-2) \cdot \sum (y_i - \bar{y})^2}}} \quad (8.7)$$

Полученное по формуле (8.7) значение сопоставляется с критическими значениями критерия Стьюдента для числа степеней свободы  $n' = n - 2$ , и делается вывод о статистической значимости полученных коэффициентов регрессии, т.е. о вероятности ошибки предсказания.

### **Множественная линейная регрессия**

Аналогично рассмотренному выше случаю, когда значение переменной  $Y$  определяется по данным значениям только одной переменной  $X$ , можно построить функцию предсказания, когда значение зависимой переменной  $Y$  определяется по значениям нескольких предикторов (8.8).

$$Y = F(X_1, X_2, \dots) \quad (8.8)$$

Например, можно предсказывать успеваемость студентов только по IQ, измеренному в начале года, а можно к IQ добавить другие независимые переменные, такие как личностные особенности, социальный статус и т.д.

В общем случае уравнение множественной линейной регрессии имеет вид (8.9).

$$Y = b_0 + b_1 \cdot X_1 + b_2 \cdot X_2 + \dots \quad (8.9)$$

В заключение приведём пример использования множественной линейной регрессии для оценки профессионально-личностного потенциала. Р. Кеттеллом были получены профессиональные портреты эффективных профессионалов следующих категорий [61, с. 174]: «продавцы», «полицейские», «психотерапевты» и «психологи-исследователи» (табл. 8.2). В качестве предикторов здесь используются факторы опросника Р.Кеттелла 16PF.

**Таблица 8.2 Функции предсказания эффективности профессиональной деятельности**

Профессия	Функция предсказания эффективности профессиональной деятельности
Продавец	$Y = 0,44A - 0,33L + 0,44Q2 + 0,22H - 0,22E - 0,22Q4$
Психотерапевт	$Y = 0,72A + 0,29B + 0,29H + 0,29N$
Психолог-исследователь	$Y = 0,31A + 0,78B + 0,47N$
Полицейский	$Y = -0,47A - 0,35F - 0,35I + 0,23Q2 + 0,23Q3$

### **Контрольные вопросы и задания**

1. Вычислите значения коэффициентов регрессии для следующих случаев: сильная положительная линейная связь ( $r=1$ ), сильная отрицательная линейная связь ( $r=-1$ ), отсутствие линейной связи ( $r=0$ ).
2. Вычислите значения коэффициентов регрессии для примера, рассмотренного в параграфе.
3. Дайте психологическую интерпретацию профессиональных портретов представителей профессий, исходя из содержания факторов 16PF Р.Кеттелла (табл. 8.2).

## 8.2. Дисперсионный анализ

Дисперсионный анализ (*ANOVA – Analysis of Variance*) – аналитико-статистический метод изучения влияния отдельных переменных, а также их сочетаний на изменчивость изучаемого признака. Метод основан на разложении общей дисперсии на составляющие компоненты, сравнивая которые можно определить долю общей вариации результирующего признака, обусловленную действием на него как регулируемых, так и неучтённых факторов. Так, профессиональная успешность в каком-либо виде деятельности (результирующий признак) обусловлена действием таких факторов, как уровень образования, опыт профессиональной деятельности, возраст, отдельные психологические особенности и т.д., а также влиянием ряда неучтённых факторов.

При осуществлении дисперсионного анализа результаты наблюдения группируются с учётом градаций каждого действующего фактора. Если фактор оказывает влияние на результирующий признак (профессиональную успешность в нашем случае), то средние результирующего признака изменяются в соответствии с градациями фактора. Внутри каждой группы (по градациям выбранного фактора) существует своя дисперсия распределения результирующего признака. Нулевая гипотеза в дисперсионном анализе сводится к предположению о равенстве внутригрупповых средних (никакого систематического действия факторов на результирующий признак нет, наблюдаемые различия в групповых средних случайны). В дисперсионном анализе используется F-критерий Фишера, позволяющий ответить на вопрос: не слишком ли велика дисперсия средних, полученных в данном исследовании, чтобы можно было утверждать, что все они являются средними случайных выборок из одной и той же генеральной совокупности. Если нулевая гипотеза будет отвергнута, то это означает, что независимые переменные в действительности формируют различные совокупности наблюдений, для которых и были получены эти выборочные средние.



Дисперсионный анализ позволяет учитывать не только влияние каждого фактора на результирующий признак по отдельности, но и совместное действие факторов во всех возможных сочетаниях. Действие неучтённых факторов учитывается суммарно.

**Рассмотрим пример.** Исследователь хочет выяснить, отличаются ли по эффективности четыре метода изучения темы «Дисперсионный анализ» студентами-психологами. Для этого планируется следующий эксперимент:

- 10 человек изучают дисперсионный анализ на лекциях.
- 10 человек читают соответствующий раздел учебника.
- 10 человек получают задание найти все опечатки в формулах в этом разделе.
- 10 человек изучают дисперсионный анализ посредством справочной системы программы «Statistica».

Четыре варианта условий предполагают различную активность студентов при изучении новой темы. Таким образом, исследователь хочет знать, будут ли различные значения фактора «активность студента» определять различную успеваемость, определяемую по 100-балльной шкале по результатам контрольной работы.

Результирующую успеваемость можно представить в виде матрицы (8.10), столбцы которой содержат результаты студентов каждой из 4 групп в соответствии с условиями эксперимента:

$$\begin{array}{cccc}
 X_{11} & X_{12} & X_{13} & X_{14} \\
 X_{21} & X_{22} & X_{23} & X_{24} \\
 \dots & \dots & \dots & \dots \\
 X_{10.1} & X_{10.2} & X_{10.3} & X_{10.4}
 \end{array} \quad (8.10)$$

где  $X_{ij}$  – успеваемость  $i$ -го студента ( $i= 1, \dots, 10$ ) из  $j$ -й подгруппы ( $j=1, \dots, 4$ );

В общем случае число групп –  $N_j$ , а число испытуемых в каждой группе –  $N_i$ . Общее число наблюдений  $N= N_i * N_j$ .

Обозначим

$M$  – среднее значение в объединённой выборке,

$M_j$  ( $j=1 \dots 4$ ) – среднее значение в соответствующей группе.

$$M = \frac{\sum_{j=1}^4 \sum_{i=1}^{10} X_{ij}}{40} \quad (8.11)$$

$$M_j = \frac{\sum_{i=1}^{10} X_{ij}}{10} \quad (8.12)$$

в общем случае

$$M = \frac{\sum_{j=1}^{N_j} \sum_{i=1}^{N_i} X_{ij}}{N} \quad (8.13)$$

Тогда

$$X_{ij} = M + \alpha_j + e_{ij}, \quad (8.14)$$

где

$X_{ij}$  –  $i$ -е значение в  $j$ -й группе,

$\alpha_j$  – константа для всех 10 значений в группе  $j$ , характеризующая изменение значений, полученных в одинаковых условиях.

$$M_j = M + \alpha_j \quad (8.15)$$

$e_{ij}$  – отклонение от группового среднего для каждого значения группы.

Заметим, что формулу (8.14) можно преобразовать следующим образом (8.16):

$$X_{ij} = M + (M_j - M) + (X_{ij} - M_j) \quad (8.16)$$

Выборочная дисперсия для каждой группы будет равна

$$s_j^2 = \sum_1^{10} \frac{(X_{ij} - M_j)^2}{9} \quad (8.17)$$

в общем случае

$$s_j^2 = \sum_1^{N_j} \frac{(X_{ij} - M_j)^2}{(N_j - 1)} \quad (8.18)$$

При разложении дисперсии в дисперсионном анализе используется понятие «сумма квадратов» (sum of squares), обозначается SS. Для вычисления этой величины воспользуемся уравнением (8.16), вычтя предварительно из обеих частей общее среднее  $M$ .

$$X_{ij} - M = (M_j - M) + (X_{ij} - M_j) \quad (8.19)$$

Возведём в квадрат и просуммируем обе части выражения (8.19) по  $i$  и по  $j$ .

$$\sum_{j=1}^4 \sum_{i=1}^{10} (X_{ij} - M)^2 = \sum_{j=1}^4 \sum_{i=1}^{10} ((M_j - M) + (X_{ij} - M_j))^2 \quad (8.20)$$

Преобразовав выражение (8.20) с помощью формулы бинома Ньютона, получим:

$$\sum_{j=1}^4 \sum_{i=1}^{10} (X_{ij} - M)^2 = \sum_{j=1}^4 \sum_{i=1}^{10} (M_j - M)^2 + \sum_{j=1}^4 \sum_{i=1}^{10} (X_{ij} - M_j)^2 \quad (8.21)$$

В выражении (8.21) обозначим

$$\sum_{j=1}^4 \sum_{i=1}^{10} (X_{ij} - M)^2 = SS_t \quad (8.22)$$

– общая сумма квадратов вариаций (total – общий).

$$\sum_{j=1}^4 \sum_{i=1}^{10} (M_j - M)^2 = 10 \sum_{j=1}^4 (M_j - M)^2 = SS_b \quad (8.23)$$

– сумма квадратов вариаций между группами (between – между)

$$\sum_{j=1}^4 \sum_{i=1}^{10} (X_{ij} - M_j)^2 = SS_w \quad (8.24)$$

– сумма квадратов вариаций внутри групп (within – внутри).

Таким образом, полная вариация разложена на 2 аддитивные компоненты, характеризующие вариации между группами и вариации внутри каждой группы (8.25).

$$SS_t = SS_b + SS_w \quad (8.25)$$

### Степени свободы

Для совокупности наблюдений число степеней свободы совпадает с числом независимых друг от друга наблюдений. Допустим,

известно, что сумма десяти чисел равна 27. Первые 9 из 10 чисел могут быть любыми, например 30, 35, 40, 45, 50, 55, 60, 65, 70, что в сумме даст 450. Это означает, что **последнее, десятое число обязательно должно** равняться – 423 и уже не является независимым. Следовательно, в данном примере имеется 9 степеней свободы.

### Критерий Фишера

Вернёмся к примеру об эффективности метода преподавания. В данном случае число степеней свободы определяется так:

$$\text{Между группами } df_b = N_j - 1 = 4 - 1 = 3 \quad (8.26)$$

$$\text{Внутри каждой группы } df_w = N_i - 1 = 10 - 1 = 9 \quad (8.27)$$

$$\text{Полное число степеней свободы } df_t = N_i \cdot N_j - 1 \quad (8.28)$$

Путём соотнесения сумм квадратов отклонений к числу степеней свободы получаются выборочные дисперсии:

а) общая по комплексу

$$s_t^2 = \frac{SS_t}{df_t} \quad (8.29)$$

б) межгрупповая или факторная

$$s_b^2 = \frac{SS_b}{df_b} \quad (8.30)$$

в) внутригрупповая или остаточная

$$s_w^2 = \frac{SS_w}{df_w} \quad (8.31)$$

Критерием оценки влияния на признак регулируемых в эксперименте факторов служит F-критерий Фишера, который вычисляется как отношение межгрупповой и внутригрупповой дисперсий:

$$F = \frac{s_b^2}{s_w^2} \quad (8.32)$$

При выполнении условия

$$F_{\text{эмп}} \geq F_{\text{кр}} \quad (8.33)$$

нулевая гипотеза отклоняется на соответствующем уровне статистической значимости.

### Двухфакторный дисперсионный анализ (общий подход)

Данные в двухфакторном дисперсионном анализе классифицируются по двум факторам. Например, нам нужно сравнить эффективность воздействия трёх различных методов обучения (фактор А) на студентов различного пола (фактор В). Фактор А имеет 3 градации, а фактор В – 2 градации. Возможны 6 различных комбинаций уровней факторов. Каждое наблюдение в двухфакторном дисперсионном анализе обозначается переменной с тремя индексами:  $X_{ijk}$ , где  $i$  – индекс фактора А, принимающий значения 1,2,3;  $j$  – индекс фактора В, принимающий значения 1, 2;  $k$  – индекс, характеризующий наблюдение внутри каждой ячейки, принимающий в нашем примере значения 1,2,3,4 (по 4 испытуемых в каждой ячейке). Сгруппированные данные представлены в таблице 8.3. В общем случае важно выполнение требования: каждой градации фактора А соответствует одинаковое количество градаций фактора В.

**Таблица 8.3. Данные, сгруппированные по факторам А и В**

Фактор А – метод обучения	Фактор В – пол	
	Юноши	Девушки
1	$X_{111}$	$X_{121}$
	$X_{112}$	$X_{122}$
	$X_{113}$	$X_{123}$
	$X_{114}$	$X_{124}$
2	$X_{211}$	$X_{221}$
	$X_{212}$	$X_{222}$
	$X_{213}$	$X_{223}$
	$X_{214}$	$X_{224}$
3	$X_{311}$	$X_{321}$
	$X_{312}$	$X_{322}$
	$X_{313}$	$X_{323}$
	$X_{314}$	$X_{324}$

Суть двухфакторного дисперсионного анализа заключается в разложении вариации значений признака на слагаемые:

$$X_{ijk} = M + \alpha_i + \beta_j + \alpha\beta_{ij} + e_{ijk}, \quad (8.34)$$

где

$M$  – среднее (одинаково для всех значений)

$\alpha_i$  – поправка для каждой градации фактора А.

$\beta_j$  – поправка для каждой градации фактора В.

$\alpha\beta_{ij}$  – поправка, отражающая факт взаимодействия факторов А и В.

$e_{ijk}$  – поправка, характеризующая различия внутри каждой ячейки.

Соответственно, сумма квадратов раскладывается на следующие компоненты:

$$\sum_{i=1}^3 \sum_{j=1}^2 \sum_{k=1}^4 (X_{ijk} - M)^2 = SS_A + SS_B + SS_{AB} + SS_w \quad (8.35)$$

Гипотезы в двухфакторном дисперсионном анализе формулируются и проверяются отдельно для проверки влияния каждого из факторов А и В независимо друг от друга, а также для проверки влияния взаимодействия этих факторов.

#### **Статистические гипотезы:**

##### *Фактор А.*

$H_0$ : Различия в эффективности обучения, обусловленные методом обучения, являются не более выраженными, чем случайные различия в эффективности обучения студентов.

$H_1$ : Различия в эффективности обучения, обусловленные методом обучения, являются более выраженными, чем случайные различия в эффективности обучения студентов.

##### *Фактор В.*

$H_0$ : Гендерные различия в эффективности обучения не более выражены, чем случайные различия между студентами.

$H_1$ : Гендерные различия в эффективности обучения более выражены, чем случайные различия между студентами.

##### *Взаимодействие факторов А и В.*

$H_0$ : Влияние метода обучения на результат одинаково и для юношей и для девушек.



$H_1$ : Существуют гендерные различия влияния метода обучения на результат.

Вычисления критерия Фишера аналогичны. Число степеней свободы

$$\begin{aligned}df_A &= I - 1 = 3 - 1 = 2 \\df_B &= J - 1 = 2 - 1 = 1 \\df_{AB} &= (I - 1)(J - 1) = 2 \\df_w &= IJ(n - 1) = 3 \cdot 2 \cdot 3 = 18\end{aligned}\tag{8.36}$$

### Контрольные вопросы и задания

1. Изложите в общем виде модель однофакторного дисперсионного анализа для несвязных выборок.
2. Сформулируйте в общем виде нулевую гипотезу в однофакторном дисперсионном анализе.
3. Что такое суммы квадратов?
4. Каковы основные ограничения применения ANOVA?
5. Каким образом учитывается действие различных неучтённых факторов в дисперсионном анализе?

### 8.3. Группировка испытуемых методами дискриминантного и кластерного анализа

Дискриминантный анализ относится к числу методов многомерного статистического анализа данных. Цель дискриминантного анализа в психологических исследованиях состоит в том, чтобы на основе имеющихся значений ряда индивидуальных показателей (характеристик) испытуемого отнести его к одной из нескольких классификационных групп некоторым оптимальным способом. Фактически дискриминантный анализ представляет собой нахождение функции предсказания (см. регрессионный анализ) для случая, когда зависимая переменная измеряется по номинативной шкале. Исходные данные для дискриминативного анализа представляются в виде традиционной таблицы данных, где строки (регистры) соответствуют испытуемым (объектам дискриминантного

анализа), число строк равно  $N$  – объёму выборки, а столбцы – измеренным показателям, число столбцов равно  $k$  – числу измеренных показателей  $T_1, T_2, \dots, T_k$  (табл. 8.4)

**Таблица 8.4. Таблица исходных данных для дискриминантного анализа**

№ испытуемого	Измеренные показатели			
	$T_1$	$T_2$	...	$T_k$
1	$X_{11}$	$X_{12}$		$X_{1k}$
2	$X_{21}$	$X_{22}$		$X_{2k}$
...	...	...		...
$N$	$X_{N1}$	$X_{N2}$	...	$X_{Nk}$

В ходе дискриминантного анализа решается задача определения решающего правила, позволяющая по значениям показателей  $T_1, T_2, \dots, T_k$  отнести конкретного испытуемого к одной из нескольких классификационных групп (которые представляют собой переменную  $Y$ , измеренную по номинативной шкале). Также определяется «вес» (значимость) каждого исходного показателя  $T_1, T_2, \dots, T_k$  для разделения объектов на классы.

Рассмотрим два примера таких задач.

**Пример 1:** Определение типа темперамента по Г.Ю.Айзенку. На основании значений показателей интроверсии и нейротизма испытуемого относят к одной из четырёх групп: сангвиники, холерики, флегматики, меланхолики.

**Пример 2:** Использование дискриминантного анализа в психотерапии для улучшения диагностики и лечения. В качестве объектов выступают криенты психотерапевта, сгруппированные по эффекту оказанной помощи. Переменными являются симптомы заболевания, социальные и психологические показатели, а также характеристики видов психотерапевтической помощи. Применение дискриминантного анализа поможет определить переменные, наиболее существенные для эффекта психотерапии, а также в дальнейшем предсказать результативность каждого вида психотерапии для конкретного клиента [55, с. 283].

Отличие кластерного анализа от дискриминантного заключается в том, что в дискриминантном анализе число групп известно заранее (например, четыре типа темперамента), тогда как в кластерном анализе число групп (кластеров) определяется в результате анализа.

Под кластером понимается группа объектов, обладающих свойством плотности (плотность объектов внутри кластера выше, чем вне его) и отделимостью от других факторов. Классическим примером кластеров являются созвездия.

Методы кластеризации делятся на агломеративные, или объединительные (агломерат – скопление) и итеративные дивизивные (division – разделение).

В ходе объединительного кластерного анализа происходит последовательное объединение наиболее близких объектов в один кластер. Процесс такого последовательного объединения иллюстрируется с помощью графика в виде дендрограммы (дерева объединения).

Различие в подходах в кластерном анализе связано с целью, для которых он проводится. По определению Д.Краскела [30, с. 24] существует следующие типы целей:

- исследовательская («просто посмотреть, что тут такое»);
- облегчить сравнение и оценку данных;
- подготовить проведение последующего анализа: после процедуры кластерного анализа можно
  - а) перед проведением более детального обследования уменьшить исходную выборку испытуемых, включив в неё представителей различных кластеров;
  - б) провести дополнительное аналитическое исследование (например, регрессионный анализ) внутри каждого кластера. Если свойства объектов из разных кластеров действительно различны, то вполне возможно, что при переходе от одного кластера к другому коэффициенты регрессии могут изменяться в широких пределах. Пример: предсказание успешности профессиональной деятельности на основе предварительного анализа интел-

лектуальных и личностных особенностей у представителей различных кластеров.

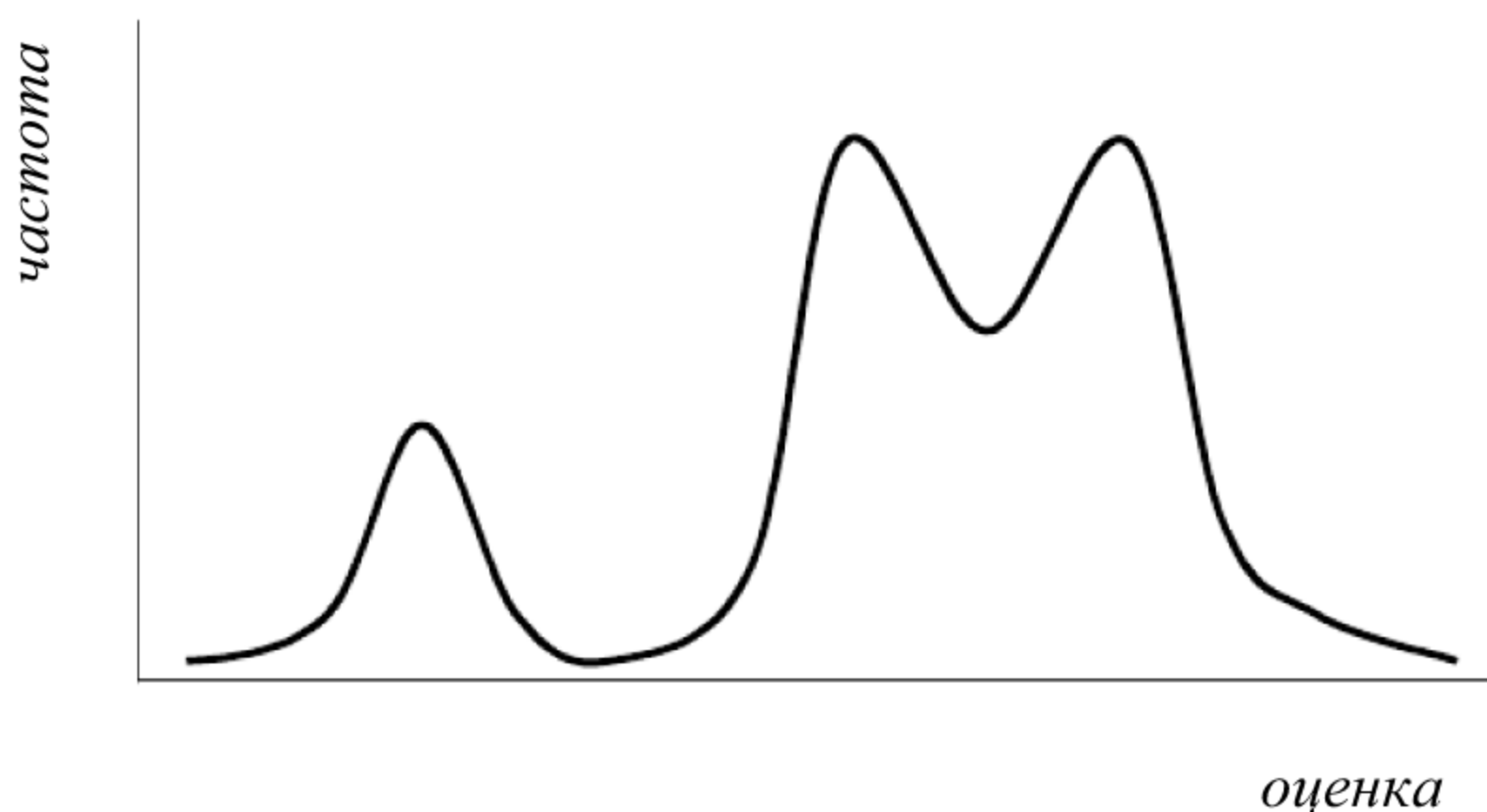
- кластерный анализ ради него самого (построение генеалогических деревьев для языков, растений или животных).

Теоретически кластерного анализа может применяться как для определения естественных кластеров (именно в этом качестве он используется в психологических исследованиях), так и для произвольной классификации объектов (пример – разбиение города на избирательные округа). Вторым вариантом кластерного анализа также может быть использован для вышеупомянутых целей (например, исследование школьников различных классов одной параллели).

Для методов кластерного анализа, отыскивающих естественные кластеры, существенен выбор критерия, с помощью которого эти кластеры выделяются. Таким критерием может быть:

- компактность кластеров;
- чёткость их разделения.

**Рассмотрим пример:** при исследовании выборки испытуемых получено распределение частот, представленное на рис. 8.2.



**Рис. 8.2** Распределение частоты оценок в выборке

Если кластеры выделяются по критерию чёткости разделения, то эти данные разумно разделить на два кластера по самой глубокой впадине. Если же выбран критерий компактности кластеров, то данные делятся на три кластера.

Исходными данными для анализа данных психологических исследований является матрица параметров объектов (таблица, в которой каждому испытуемому соответствует ряд показателей, т.е. результатов выполнения тестов, экспертных оценок и т.д.).

На следующем этапе выбирается мера сходства объектов. Правильный выбор меры сходства и способа вычисления расстояния между испытуемыми – один из самых важных моментов.

В программе STATISTICA используются следующие меры сходства объектов:

- евклидова метрика (сумма квадратов разностей по каждому параметру);
- манхэттенское расстояние (сумма по координатных расстояний);
- коэффициент корреляции Пирсона;
- коэффициент ассоциации (совстречаемости).

Помимо мер сходства используются и различные методы кластеризации, наиболее часто используемым является метод К-средних.

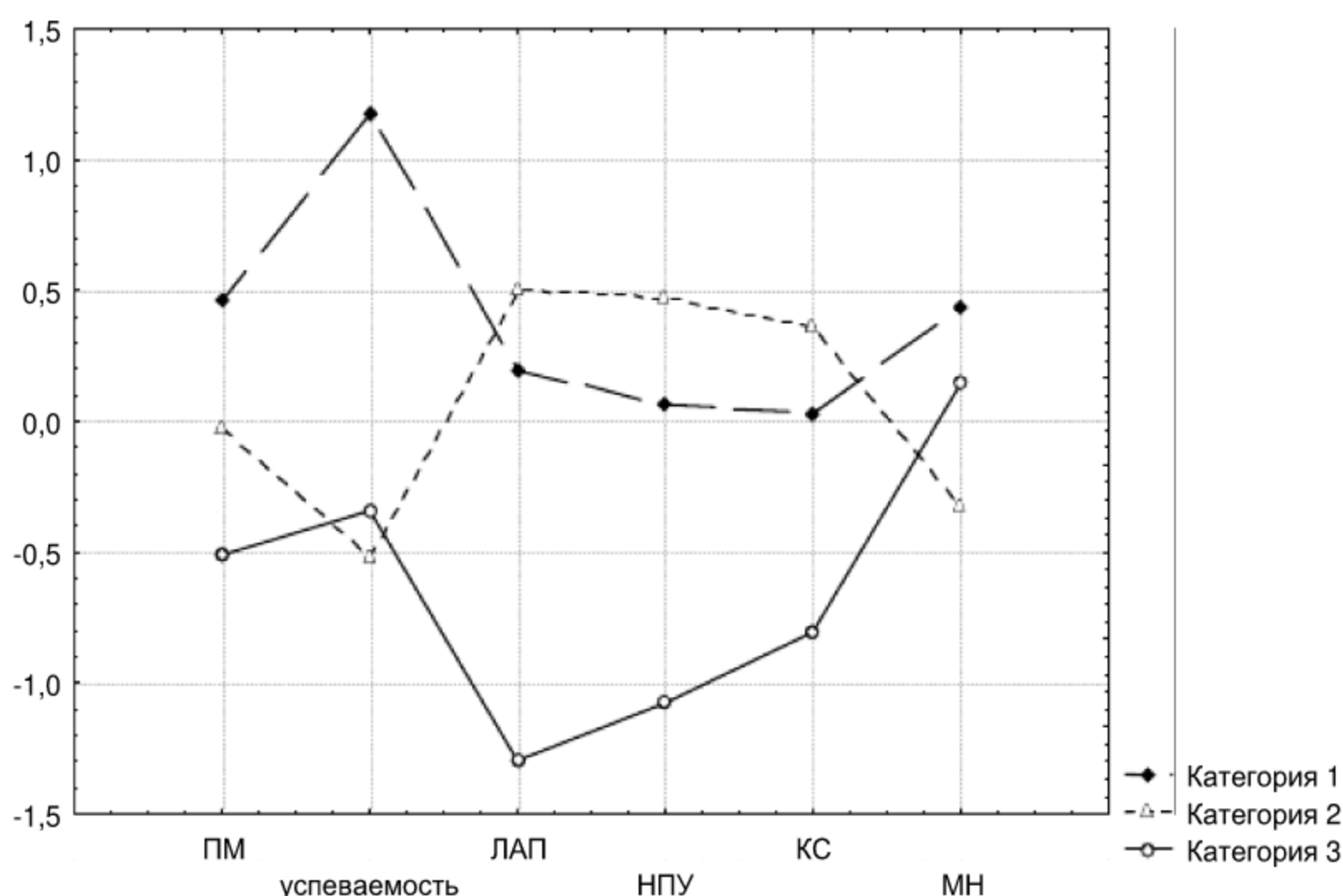
Перед началом классификации данные необходимо привести к стандартному виду (вычесть среднее и разделить на стандартное отклонение). Иногда имеет смысл уменьшить количество признаков.

После того, как классификация проведена, необходимо сопоставить полученные кластеры попарно по каждому значимому показателю с помощью статистических критериев различий, и дать психологическое описание каждому кластеру.

**Пример.** В исследовании С.Р.Хамидуллиной изучались особенности адаптации первокурсников к обучению в вузе. Изначально предполагалось разделить выборку на адаптировавшихся и неадаптировавшихся студентов и проанализировать степень выраженно-

сти у каждой категории студентов таких показателей, как НПУ (нервно-психическая устойчивость), КС (коммуникативные способности), МН (моральная нормативность), ПМ (самооценка психологического микроклимата в группе), успеваемость (по результатам 1 сессии). Всего было обследовано 78 чел.

Для разделения выборки использовался итеративный дивизивный метод кластеризации, в ходе которого выяснилось, что испытуемые разделились не на две, а на три категории (рис. 8.3).



**Рис. 8.3 Результаты кластерного анализа выборки первокурсников (данные стандартизированы)**

*Первая категория (28%)* – студенты с хорошей успеваемостью и высоким показателем моральной нормативности. Нервно-психическая устойчивость и коммуникативные способности у них на среднем уровне. Самооценка психологического микроклимата в учебной группе высокая. Можно сказать, что эта категория студентов полностью адаптировалась к обучению в вузе.

*Вторая и третья категории* – студенты с низкой успеваемостью, но остальные показатели социально-психологической адаптации у них существенно различаются.



*Вторая категория (49%)* отличается средней самооценкой психологического микроклимата в учебной группе, низкими значениями показателя моральной нормативности и высокими значениями остальных показателей адаптации (интегративный показатель адаптивности ЛАП у них даже несколько выше, чем у первой категории студентов). Мы считаем, что второй категории студентов по тем или иным причинам не удастся демонстрировать высокую успеваемость, но они смирились со своим «учебным статусом» и чувствуют себя вполне комфортно. Условно эту группу студентов можно назвать «разгильдяями» (похожие показатели могут быть и у работающих студентов, но на первом курсе студенты, как правило, еще не работают).

*Третья категория (23%)* – студенты, которые не смогли адаптироваться к обучению в высшей школе. Психологический микроклимат в своей учебной группе они оценивают как низкий, у них также низкие показатели нервно-психической устойчивости, коммуникативных способностей и в целом низкий адаптивный потенциал (ЛАП). Однако у них достаточно высокие оценки по моральной нормативности. Из этого следует, что данная категория студентов не удовлетворена своим положением, переживает из-за сложившейся ситуации. Именно этой категории студентов в первую очередь требуется психологическая поддержка со стороны как преподавателей, так и одногруппников.

Таким образом, группа студентов с низкой успеваемостью неоднородна. Административные меры воздействия могут оказаться действенными только для второй категории студентов и усугубить ситуацию для третьей категории. Социально-психологической адаптации (и как следствие – повышения успеваемости) студентов третьей категории могут способствовать социально-психологические тренинги, психотехнические и деловые игры, развивающие коммуникативные способности и улучшающие психологический микроклимат в студенческих группах, а также организация психологического консультирования.

### **Контрольные вопросы и задания**

1. В чём заключается цель применения дискриминантного анализа при обработке эмпирических данных?
2. В чём основное отличие дискриминантного анализа от кластерного? Перечислите основные типы целей применения кластерного анализа при обработке данных психологических обследований.
3. Что такое кластер? На какие две группы делятся методы кластеризации? Как называется график, иллюстрирующий процесс последовательного объединения объектов в кластеры?
4. Какие критерии могут использоваться для выделения кластеров из совокупности наблюдений?
5. Какие меры сходства объектов Вам известны? Охарактеризуйте суть евклидовой метрики и коэффициента корреляции Пирсона.

#### **8.4. Группировка переменных с помощью методов факторного анализа**

Существует множество психологических исследований, целью которых является идентификация психических черт (это могут быть черты личности, способности и т.д.) Такие исследования начинаются с вычисления интеркорреляций (взаимных корреляций) показателей, полученных на выборке испытуемых по широкому набору психодиагностических методик. Значения полученных коэффициентов корреляции составляют матрицу интеркорреляций, которая подвергается математической обработке с целью выявления общих факторов. Полученные факторы и определяют искомые психические черты.

Основной целью факторного анализа является сокращение размерности пространства описания данных, т.е. числа переменных (**редукция данных**) и определение структуры взаимосвязей между переменными (**структурная классификация переменных**).

Факторный анализ не используется для обработки первичных экспериментальных данных, полученных непосредственно при обследовании испытуемых. Исходными данными для факторного анализа являются корреляционные связи между переменными. Как правило, факторный анализ осуществляется на основе коэффи-

ентов корреляции Пирсона, но его можно осуществить и для коэффициентов ранговой корреляции [92, с. 375].

Итак, факторному анализу подвергаются корреляционные матрицы (матрицы интеркорреляций). Главное понятие факторного анализа – фактор. Это искусственный статистический показатель, возникающий в результате преобразований матрицы интеркорреляций. Процедура извлечения факторов из матрицы интеркорреляций называется факторизацией матрицы. В результате факторизации получается факторная матрица, т.е. таблица, показывающая вес или нагрузку каждого из факторов по каждому показателю. Факторные нагрузки, являющиеся элементами факторной матрицы представляют собой коэффициенты корреляции данного фактора со всеми исходными показателями.

#### **Условия применения факторного анализа.**

- Все данные должны быть количественно измерены по шкале порядка и выше.
- Все переменные должны быть независимыми, а их распределение должно приближаться к нормальному (допускается отличие от нормальности на уровне  $p=0,1$ ).
- Связи между переменными не должны иметь явно криволинейного характера.
- В исходной матрице интеркорреляций должно быть несколько значений коэффициента корреляции по модулю выше 0,3. Иначе процесс извлечения факторов бесперспективен.
- Определение переменных, принадлежащих тому или иному выделенному фактору, производится по абсолютной величине нагрузки по данному фактору. Минимальный уровень величины нагрузки равен (по модулю) 0,4, но во многих случаях достигает 0,7.
- Теоретически объём выборки испытуемых должен быть не менее 100 человек. Однако, его можно несколько уменьшить, если имеющиеся данные получены с помощью валидных тестов и нет сомнений в их надёжности, а факторы определились с нагрузками не меньше 0,7.

**Рассмотрим пример** А.Анастази [4, с. 334]. Выборка испытуемых была обследована с помощью батареи из 10 тестов. После факторизации матрицы интеркорреляций была получена факторная матрица, представленная в таблице 8.4.

**Таблица 8.4. Факторная матрица**

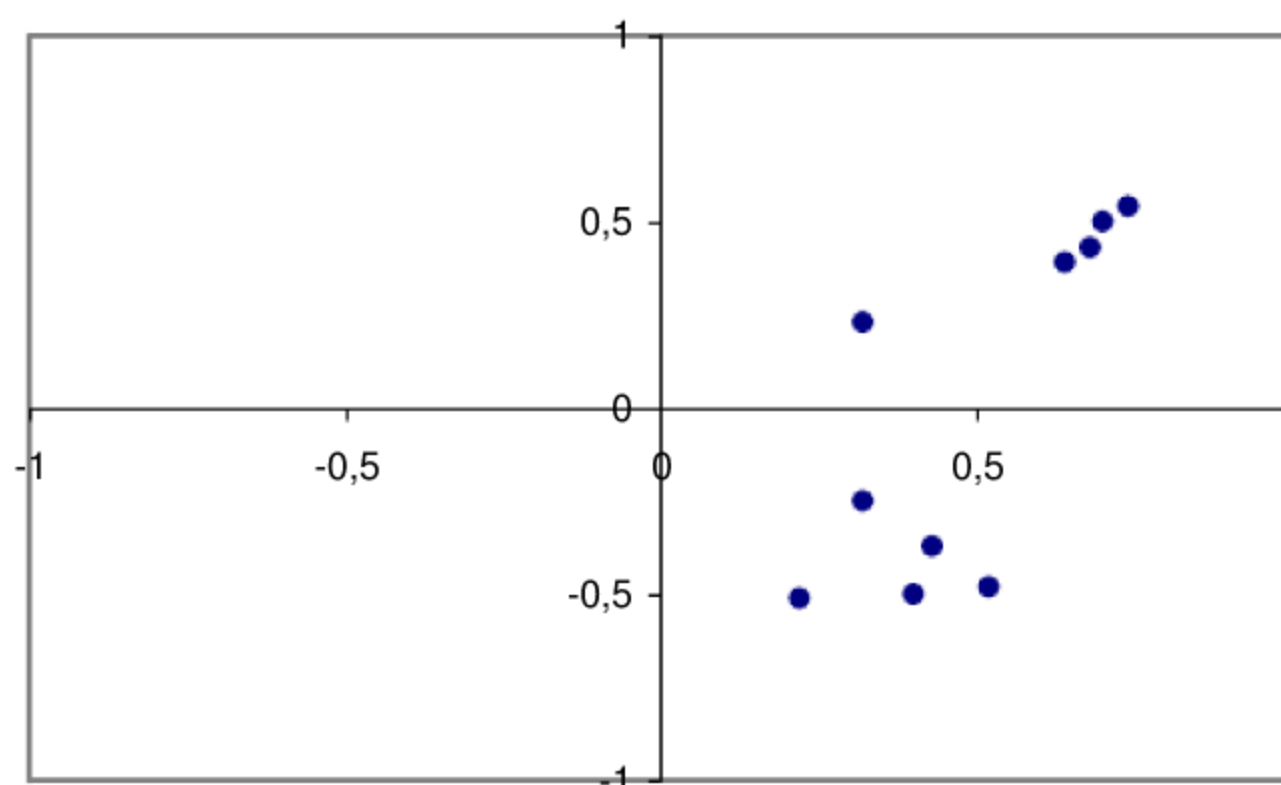
Тест	Фактор 1	Фактор 2
Словарный	0,74	0,54
Аналогии	0,64	0,39
Завершение предложений	0,68	0,43
Восстановление порядка слов в предложении	0,32	0,23
Понимание прочитанного	0,7	0,5
Сложение	0,22	-0,51
Умножение	0,4	-0,5
Арифметические задачи	0,52	-0,48
Составление уравнений	0,43	-0,37
Завершение числовых рядов	0,32	-0,25

В факторном анализе принято представлять факторы геометрически в виде осей координат, относительно которых каждый тест может быть изображён в виде точки. Координаты каждой точки (т.е. теста) определяются величиной нагрузки каждого фактора по этому тесту. В результате получится своеобразная диаграмма рассеяния (рис. 8.4).

На диаграмме рассеяния видно, что 10 тестов могут быть сгруппированы в два кластера (тесты с 1 по 5 образуют кластер в верхней части координатной плоскости, а тесты с 6 по 10 образуют кластер в нижней части координатной плоскости).

Заметим, что положение осей координат не фиксировано данными, так как исходная таблица интеркорреляций определяет лишь положение тестов друг относительно друга. Те же точки можно нанести на плоскость с любым положением координатных осей. По этой причине при проведении факторного анализа обычно вращают

оси до тех пор, пока не получается наиболее приемлемое и легко интерпретируемое отображение.



**Рис. 8.4 Графическое представление факторной матрицы**

Критериями приемлемости (по Л.Терстоуну) считаются:

- положительное многообразие;
- простая структура.

**Положительное многообразие** предполагает вращение осей до положения, при котором исключаются все значимые отрицательные веса. Критерий **простой структуры** означает, что каждый тест должен иметь нагрузки по как можно меньшему числу факторов. Так, если тест имеет высокую нагрузку по одному фактору и не имеет нагрузок по другим факторам, то природу этого фактора определить достаточно легко. Если же тест имеет средние нагрузки по нескольким факторам, то о природе любого из них судить гораздо сложнее.

Существует понятие независимости факторов. Факторы являются независимыми, если корреляция между ними (исходя из нагрузок факторов в факторной матрице) равна нулю. Такие факторы называются ортогональными. В противном случае факторы считаются зависимыми (облическими). Ортогональные факторы в отли-

чие от облических дают более простые варианты взаимодействия внутри факторной матрицы.

Иллюстрацией ортогональных факторов является задача Л.Терстоуна, который, взяв ряд коробок разных размеров и формы, измерил в каждой из них более 20 различных показателей и вычислил корреляции между ними. Подвергнув полученную матрицу интеркорреляций факторному анализу, он получил три фактора, корреляция между которыми была равна нулю. Этими факторами были «длина», «ширина» и «высота» [22, с.276].

В практике факторного анализа используются разные варианты вращения факторов, при этом выделяется два основных метода вращения: ортогональное и косоугольное (облическое). Сущность ортогонального вращения заключается в том, что при вращении остаётся верным предположение о независимости факторов. Ортогональное вращение бывает четырёх видов:

- варимакс;
- кватримакс;
- эквимакс;
- бикватримакс.

При применении факторного анализа в психологии почти всегда используется метод варимакс, при котором минимизируется количество переменных, имеющих высокие нагрузки на данный фактор, при этом максимально увеличивается дисперсия фактора. Это упрощает описание фактора за счёт группировки вокруг него только тех переменных, которые связаны с ним в большей степени, чем остальные.

Дисперсия фактора определяется с помощью такой характеристики, как собственное значение фактора. Для её вычисления все нагрузки данного фактора возводятся в квадрат и складываются по столбцу. Если собственное значение фактора разделить на число переменных, то полученное число покажет, какая доля дисперсии объясняется данным фактором.

Продолжим рассмотрение примера. После вращения осей тесты 1–5 располагаются близко к оси 1', а тесты 6–10 тесно группиру-



ются вокруг оси 2'. Результирующая факторная матрица принимает вид (табл. 8.5).

Такого рода факторная матрица уже может служить основой для интерпретации и наименования факторов. Этот этап уже не связан со статистикой, а требует психологической подготовки и интуиции. Для понимания природы каждого фактора необходимо изучить тесты, имеющие высокие нагрузки по этому фактору и попытаться обнаружить в них общие психологические процессы. Чем больше оказывается тестов с высокой нагрузкой по фактору, тем легче раскрыть его природу.

**Таблица 8.5. Факторная матрица после вращения факторов**

Тест	Фактор 1	Фактор 2
Словарный	0,91	-0,06
Аналогии	0,75	0,02
Завершение предложений	0,8	0,00
Восстановление порядка слов в предложении	0,39	0,02
Понимание прочитанного	0,86	-0,04
Сложение	-0,09	0,55
Умножение	0,07	0,64
Арифметические задачи	0,18	0,68
Составление уравнений	0,16	0,54
Завершение числовых рядов.	0,13	0,38

В рассмотренном примере А.Анастази фактор 1 соответствует вербальным способностям, а фактор 2 – арифметическим способностям. Таким образом, многообразие показателей было сведено к двум обобщённым показателям, на основании которых можно дать более лаконичную характеристику испытуемым. Безусловно, подобная лаконичность влечёт за собой потерю в точности характеристик. Нахождение разумного баланса между обилием показателей и потерей точности – один из показателей квалификации исследователя.

### **Контрольные вопросы и задания**

1. В каких случаях применяется факторный анализ? Что является исходными данными для факторного анализа? Какова основная цель факторного анализа?
2. Дайте определение следующим понятиям: фактор, факторизация матрицы интеркорреляций, факторная матрица, факторные нагрузки, ортогональные факторы, облические факторы.
3. Объясните целесообразность процедуры вращения факторов. Что включают в себя критерии положительного многообразия и простой структуры?

## **Глава 9.**

### **Использование методов измерения и оценки в психологических исследованиях**

#### **9.1. Правила составления пакета методик психологического исследования.**

Выбор материала из всего арсенала методик – очень важный вопрос. От его решения во многом зависят надёжность и достоверность результатов исследования. А.Анастази определяет следующие правила составления пакета психодиагностических методик [4]:

- Методический материал должен соответствовать целям и задачам исследования. Если перед психологом стоит задача изучения психофизических свойств индивида (например, скорости реакции или переключения внимания), то недопустимо использовать методики определения личностных особенностей.
- Необходимо строго учитывать информацию о психометрических характеристиках выбранных методик (данные о видах надёжности, валидности, трудность заданий, наличие статистических норм, данные о выборке стандартизации).
- Следует заранее оценить временные затраты на процедуру диагностики, сама процедура обследования обязательно должна быть стандартной (соблюдение всех правил регламентации: инструкция, бланки и т.д.).
- Количество методик определяется сложностью поставленных задач. Если изучается какой-либо один узкий аспект психики человека, то достаточно использовать 1-2 методики. Если же задача ставится более широко, то следует использовать пакет методик; в этом случае результаты одной методики дополняют и уточняют результаты других методик.

К настоящему моменту разработано множество психодиагностических методик, и это методический арсенал постоянно пополняется новыми разработками как отечественных, так и западных психологов, педагогов и социологов. В этих условиях существует определённая конкуренция методических средств психологического оценивания. А.Н.Поддьяков [65] рассматривает данную проблему на примере тестов способностей и познавательного развития. Учитывая исключительную важность понимания данной проблемы, приведём основные тезисы статьи А.Н.Поддьякова «Тестирование интеллекта, конкуренция и рефлексия» [66].

По его определению, понятие «конкуренция средств психодиагностики» имеет нескольких существенно разных значений.

1. Под ним может пониматься конкурентная борьба представителей различных групп за использование и распространение той или иной диагностической стратегии, того или иного диагностического средства для оценки развития обучаемых. Сюда входит доказательство преимуществ одних средств и показ недостатков других, их реклама и антиреклама, принятие решений о выборе одной диагностической стратегии и об отказе от другой и т.д. Такая борьба может вестись на различных уровнях: научном, политическом (как в случае педологии), бюрократическом, экономическом и т.д.

Помимо конкурентной борьбы средств психодиагностики между собой имеет место борьба с эффективным применением этих средств. Разрабатываются и публикуются рекомендации по оптимальным стратегиям прохождения тестирования: обходу ловушек психодиагностических методик и даже таких ответов, которые создадут желаемый психологический портрет человека, проходящего тестирование. Но в такой ситуации психодиагносты обязаны разрабатывать новые приемы, нейтрализующие «продвинутые» стратегии прохождения тестирования – хотя бы для того, чтобы обезопасить других людей от потенциально опасных субъектов [107]. Поэтому в настоящее время идет все более интенсивная гонка тестовых и противотестовых «вооружений».

2. Под конкуренцией средств диагностики можно понимать ситуации, когда два или более различных тестов, предназначенные для диагностики одних и тех же способностей, но разработанные, например, разными авторами, дают противоречащие друг другу результаты. По одному из тестов познавательного развития человек может систематически получать более высокие баллы, а по другому – значительно более низкие, и неясно, результаты какого из этих тестов считать в большей степени характеризующими данного человека. Для выбора необходимо провести специальную работу по оценке валидности теста, в том числе конкурентной валидности. Распространение или же сокрытие информации о конкурентной валидности психодиагностической методики может служить средством конкурентной борьбы в первом значении.

3. Под конкуренцией средств диагностики познавательного развития можно понимать такое их взаимодействие на внутриспсихологическом уровне, когда практическое использование одного диагностического средства снижает показатели другого из-за сложной динамики возникающих психических процессов. Одним из наиболее ярких примеров подобной интерференции тестов являются данные В.Н.Дружинина: если провести тестирование творческих способностей (креативности), а вслед за ним – тестирование интеллекта, то результаты испытуемых по тесту интеллекта оказываются ниже, чем без предварительного тестирования креативности. Разумеется, это не означает, что у испытуемых после тестирования креативности навсегда снизился интеллект. Речь идет лишь о временном снижении его тестовых показателей.

В контексте последнего тезиса отметим, что нами был проведен эксперимент, содержание которого заключалось в следующем. Две группы студентов-кибернетиков по 23 человека были обследованы с помощью методики «Стандартные прогрессивные матрицы» Д.Равена. С помощью U-критерия Манна-Уитни было доказано отсутствие значимых различий между этими группами. На втором этапе эксперимента (через 2 недели) одной из групп было предложено выполнить тест «Продвинутые прогрессивные матрицы»

Д.Равена, а второй группе предлагалось выполнить две методики: сначала тест «Закончи рисунок» Е.Торренса, а затем – тест «Продвинутые прогрессивные матрицы» Д.Равена». При этом фиксировалось время выполнения теста Равена в обеих группах. Последующий сравнительный анализ результатов и времени выполнения теста Равена в двух группах показал отсутствие статистически значимых различий в результате и времени выполнения данного теста.

Вместе с тем, нельзя не согласиться с утверждением А.Н.Поддьякова, что интерференция тестов разных типов имеет как теоретическое, так и практическое значение. Она означает относительность получаемых оценок тестирования, возможность систематических ошибок диагностики познавательного развития, а также открывает возможности преднамеренной манипуляции.

Остановимся далее на позиции А.Н.Поддьякова относительно измерения интеллектуальных, творческих и исследовательских способностей. В реальной познавательной деятельности интеллект, креативность, исследовательские способности образуют единство. Разделение этих понятий произошло с целью более детального изучения этих трёх сторон познавательной деятельности.

Вместе с тем, существует общий методологический закон исследования сложных динамических систем, к которым относится и психика. Из-за множественных положительных и отрицательных связей, существующих в сложной динамической системе, любой метод, раскрывающий одни ее характеристики, будет неизбежно скрывать, подавлять и деформировать другие ее существенные характеристики. Универсальный метод исследования, который бы только раскрывал характеристики сложной системы, ничему при этом не мешая, невозможен. **Это положение справедливо и для различных тестов познавательного развития, в ряде случаев конкурирующих между собой за наиболее адекватную репрезентацию ментального статуса тестируемого.**

Как уже отмечалось в главе 2, развитие психодиагностики фактически началось с изучения интеллектуальных способностей (Ф.Гальтон, Д.Кеттелл, А.Бине, Г.И.Россолимо, Ф.Е.Рыбаков). С



середины XX века, когда были опубликованы работы Д.Гилфорда, начали разрабатываться тесты креативности, побуждающие испытуемых к изобретению новых, разнообразных, оригинальных решений. Оказалось, что тесты интеллекта и креативности находятся в сложных и неоднозначных отношениях друг с другом. Появились различные подходы к проблеме взаимосвязи интеллекта и креативности [20].

В 60-х гг. XX века для изучения сторон познавательной деятельности, не охватываемых тестами интеллекта и креативности, стали применяться тесты исследовательского поведения [65]. Они диагностируют способности человека приобретать новую информацию при реальном взаимодействии с неизвестными объектами, действовать практически в условиях новизны и неопределенности, самостоятельно ставя и решая различные исследовательские задачи. Успешное выполнение этих тестов, как оказалось, не требует высокого интеллекта в его тестовом операциональном понимании: тесты исследовательского поведения часто имеют отрицательные корреляции с тестами интеллекта. Также были обнаружены значимые отрицательные корреляции между некоторыми показателями исследовательского поведения и творчества.

Из фактов отрицательных корреляций между различными тестами познавательных способностей вытекают очень важные практические следствия. Так, в России для оценки познавательного развития используются в абсолютном большинстве случаев только тесты интеллекта, а стандартизованных тестов исследовательского поведения нет. Это, как отмечает А.Н.Поддьяков, означает принципиальную односторонность получаемых результатов. Зачастую ребёнок, получивший низкий балл по тесту интеллекта, получает высокий балл по тесту исследовательского поведения, то есть проявил бы себя как достаточно хороший исследователь новизны и неопределенности, способный самостоятельно ставить и решать исследовательские задачи при реальном взаимодействии с объектами. К сожалению, верно и обратное: если ребенок получил высокий балл по тесту интеллекта, то с большой вероятностью он полу-

чил бы более низкий балл по тесту исследовательского поведения, и следует обратить внимание на развитие его исследовательского потенциала. Этот нюанс особенно важен в контексте перехода к новым стандартам образования, в которых результатом образования являются не только традиционные «знания, умения, навыки», но и освоенные компетенции, при этом сам стандарт написан в терминах компетенций (т.е. умения и возможности делать). Таким образом, психологическая оценка индивидуальных различий обучающихся и её учёт в образовательном процессе является одним из определяющих факторов повышения эффективности образовательного процесса.

Следующая проблема диагностики познавательных способностей названа А.Н.Поддьяковым «проблемой судейства».

Процесс выполнения теста – это социокогнитивное взаимодействие между:

а) лицами, которые создают тот или иной психодиагностический инструментарий, будучи вооружены вполне определенными теоретическими установками, познавательными и практическими целями, и находятся при этом на том или ином уровне познавательного и личностного развития;

б) лицами, исследуемыми с помощью этого инструментария, которые тоже имеют свои познавательные и практические цели и находятся на своем уровне познавательного и личностного развития – вовсе не обязательно более низком.

Это очень важный социальный аспект психодиагностики, отмечает А.Н.Поддьяков. Ведь те, кто придумывают стандарты оценки чужого интеллекта и творчества, имеют свой собственный, небезграничный интеллект и не безупречные представления о том, что такое интеллект и творчество и как их измерять. Некритическое использование результатов их работы может быть попросту опасно.

В заключение параграфа остановимся на проблеме несбалансированной сложности заданий. В контексте психометрических характеристик речь здесь идёт о содержательной валидности тестов

интеллекта. А.Н.Поддьяков приводит пример культурно обусловленных когнитивно-стилевых предпочтений, наблюдающихся при разработке тестовых заданий психологами и их выполнении испытуемыми. Было установлено (исследование Р.Нисбетта), что тесты, разработанные европейскими психологами, фактически не позволяют жителям и выходцам из Азии продемонстрировать сильные стороны своего мышления, которые менее выражены у европейцев (это произошло не намеренно, а просто потому, что европейские разработчики имели слабое представление о других возможностях). Но это означает, что для справедливого сравнения европейцам надо бы пройти тесты интеллекта, разработанные представителями азиатской культуры, иначе для первых создаются изначальные преимущества при конкуренции. Другой выход может быть в объединении высококвалифицированных разработчиков в коллективы, вырабатывающих общую точку зрения.

При этом в целом никакое конечное решение в области диагностики познавательного развития недостижимо. Человечество никогда не будет точно знать возможные направления развития творческого мышления – оно сущностно непредсказуемо. А значит, как подчёркивает А.Н.Поддьяков, человечество всегда будет не вполне готово к диагностике своего интеллектуального и творческого потенциала, ему всегда будут требоваться все новые методы выявления, оценивания и развития способностей человека – ранее неизвестных, впервые открываемых, хотя и существовавших, или же только возникающих в ходе социального развития и вступающих в сложные, подавляющие и стимулирующие отношения с уже существующими способностями.

Отметим в заключение, что подбор адекватных методов психологической оценки на фоне понимания исследователем изложенных в настоящем параграфе проблем позволит избежать многих ошибок в выводах по результатам проведённых экспериментальных исследований.

### **Контрольные вопросы и задания**

1. *Перечислите правила составления пакета психодиагностических методик, сформулированные А.Анастази. Какие ещё правила Вы бы добавили в этот перечень?*
2. *Раскройте аспекты проблемы конкуренции средств психодиагностики.*

## **9.2 Пример комплексного использования методов измерения и оценки**

В параграфе 1.1 нами были рассмотрены общие подходы к научному исследованию, изложены основные его этапы, определены роль и место методов измерения и оценки в методологии исследований в области психологии. Содержанию методов измерения и оценки были посвящены большинство глав и параграфов учебного пособия. В настоящем параграфе представлен образец комплексного использования методов измерения и оценки на примере дипломного исследования, выполненного О.С.Анохиной под нашим руководством.

### **9.2.1 Научный аппарат исследования**

**Тема** исследования: «Внутренние факторы успешности процесса социально-психологической адаптации первокурсников к обучению в вузе»

**Объект:** процесс социально-психологической адаптации первокурсников к обучению в вузе.

**Предмет:** внутренние факторы успешности процесса социально-психологической адаптации первокурсников к обучению в вузе.

**Гипотеза:** Студенты, успешно адаптировавшиеся к обучению в вузе, обладают сходным набором индивидуальных характеристик, имеющих различную выраженность в зависимости от курса обучения, при этом часть из этих характеристик сформирована у них уже к моменту поступления в вуз.

**Методы исследования:** теоретический анализ литературы, психодиагностический метод, анкетный опрос, анализ документации,

констатирующий эксперимент, методы математической статистики.

**Методики исследования:** 16-факторный личностный опросник (16ФЛО) Р. Кеттелла; многоуровневый личностный опросник «Адаптивность» (МЛО-АМ) (А.Г.Маклаков, С.В.Чермянин); методика «Изучение мотивов учебной деятельности студентов» (А.А.Реан, В.А.Якунин).

**База исследования и экспериментальная выборка:** Физико-энергетический факультет ИАТЭ НИЯУ МИФИ, студенты 1–3 курсов в количестве 101 человека.

### 9.2.2 Модель исследования

Адаптация (от лат. *adaptare* – приспособлять) – это приспособление индивида к условиям окружающей среды. **Психологическая адаптация** – это приспособление человека как личности к существованию в обществе в соответствии с требованиями этого общества и с собственными потребностями, мотивами и интересами. Основные проявления психологической адаптации – **взаимодействие** человека с окружающими людьми и его **активная деятельность**.

**Критериями успешности адаптации** студентов к обучению в вузе являются

- усвоение программы обучения;
- принятие правил и норм обучения в вузе;
- сформированность межличностных отношений с другими участниками образовательного процесса;
- эмоциональное благополучие.

В результате теоретического анализа проблемы адаптации студентов к обучению в вузе было установлено, что на успешность адаптационного процесса влияют внешние и внутренние факторы.

**Внешние:** педагогический фактор (методика преподавания, система оценки знаний); психологический фактор (психологическая поддержка преподавателей, кураторов, младших кураторов, родителей и близких друзей, психологические тренинги); социальный

фактор (проживание, питание, медицинское обслуживание, материальная поддержка родителей).

**Внутренние:** состояние здоровья; уровень школьной подготовки; учебная самостоятельность; индивидуально-психологические особенности (особенности личности, уровень интеллектуального развития, психофизиологические особенности, учебная мотивация).

В соответствии с темой исследования «Внутренние факторы успешности процесса социально-психологической адаптации первокурсников к обучению в вузе» были определены

А. Методы оценки успешности адаптации (табл. 9.1)

Б. Методы оценки внутренних факторов успешности процесса социально-психологической адаптации первокурсников (табл. 9.2)

**Таблица 9.1 Методы оценки критериев успешности адаптации**

<b>Критерии успешности адаптации</b>	<b>Методы оценки</b>
Усвоение программы обучения	Анализ документации (успеваемость)
принятие правил и норм обучения в вузе	Психодиагностический метод: МЛО-АМ, шкала МН
сформированность межличностных отношений с другими участниками образовательного процесса	Психодиагностический метод: МЛО-АМ, шкала КС
эмоциональное благополучие	Психодиагностический метод: МЛО-АМ, шкала НПУ

**Таблица 9.2 Методы оценки внутренних факторов успешности адаптации**

<b>Факторы успешности адаптации</b>	<b>Методы оценки</b>
Состояние здоровья	Анализ медицинской документации
Уровень школьной подготовки	Анализ документации (результаты ЕГЭ по математике и физике)



Учебная самостоятельность	Анкетный опрос
Психологические особенности	Психодиагностический метод:
– особенности личности	16-факторный личностный опросник Р.Кеттелла
– уровень интеллектуального развития	16ФЛО (фактор В), анализ документации (результаты ЕГЭ по математике),
– учебная мотивация	Методика «Изучение мотивов учебной деятельности студентов» (А.А. Реан, В.А. Якунин)

### 9.2.3 Методы психологической оценки

#### Многоуровневый личностный опросник «Адаптивность» (МЛО-АМ)

Опросник разработан А.Г. Маклаковым и С.В. Чермяниным [45]. В основу методики положено представление об адаптации, как о непрерывном процессе активного приспособления человека к постоянно меняющимся условиям социальной среды и профессиональной деятельности.

Опросник содержит 165 вопросов и имеет следующие шкалы:

- Шкала достоверности.
- Нервно-психическая устойчивость (НПУ). Низкие значения по данной шкале свидетельствуют о низком уровне поведенческой регуляции, определенной склонности к срывам, отсутствие адекватной самооценки. Высокие значения говорят о высоком уровне нервно-психической устойчивости и поведенческой регуляции, о высокой адекватной самооценке и реально восприятии действительности.
- Коммуникативные способности (КС). Низкие значения по данной шкале свидетельствуют о низком уровне коммуникативных способностей, о затруднениях в построении контактов с окружающими, о проявлении агрессивности, конфликтности. Высокие значения говорят о высоком уровне развития коммуника-

тивных способностей, о легком установлении контактов с окружающими.

- Моральная нормативность (МН). Низкие значения по данной шкале свидетельствуют о том, что человек не может адекватно оценить свое место и роль в коллективе, не стремится соблюдать общепринятые нормы поведения. Высокие значения свидетельствуют о том, что человек реально оценивает свою роль в коллективе, ориентируется на соблюдении общепринятых норм поведения;
- Личностный адаптационный потенциал (ЛАП). Интегративный показатель адаптации.

Исходя из описания, делаем вывод о том, что методика выбрана корректно: содержание шкал пригодно для оценки критериев адаптации (табл. 9.1).

### **16-факторный личностный опросник Р.Кеттелла (16 ФЛО)**

Опросник являет собой систему из 16 шкал, измеряющих различные полярные свойства личности. Опросник позволяет подробно описать личностную структуру, вскрыть взаимосвязь отдельных свойств личности, выявить внутриличностные проблемы, найти компенсаторные механизмы для поддержания психического здоровья. Кратко охарактеризуем факторы опросника [84].

**Фактор А** (открытость-замкнутость). Высокие значения говорят о добросердечности, эмоциональности, гибкости и легкости в общении. Низкие оценки свидетельствуют о холодности, скрытности, сдержанности и осторожности в общении.

**Фактор В** (высокий интеллект – низкий интеллект). Высокие оценки имеют люди с хорошим уровнем интеллектуального развития, с гибким, аналитичным умом. Низкие оценки свидетельствуют о плохих способностях к обучению, интеллектуальной лени, низком уровне информированности.

**Фактор С** (эмоциональная устойчивость – эмоциональная неустойчивость). Высокие оценки свойственны людям эмоционально зрелым, спокойным, уверенным в себе, реалистично смотрящим на

жизнь. Низкие оценки свидетельствуют о неумении контролировать свои эмоции, безответственности, капризности.

**Фактор Е** (доминантность – конформность). Высокие оценки по фактору имеют люди властные, независимые, стремящиеся к самостоятельности и лидерству. Низкие оценки демонстрируют покорность, зависимость, неуверенность в своих силах, неумение отстаивать свою точку зрения.

**Фактор F** (беспечность – предусмотрительность). Высокие оценки характерны для людей, легко воспринимающих жизнь, активных, бодрых, беспечных. Низкие оценки говорят об озабоченности, осторожности, медлительности, стремлении спланировать свои поступки во избежание возможных неудач.

**Фактор G** (высокая совестливость – недобросовестность). Люди с высокими показателями по этому фактору очень дисциплинированы, добросовестны, во всем любят порядок, глубоко порядочны. Низкие оценки имеют люди непостоянные, эгоистичные, неисполнительные, способные на ложь.

**Фактор H** (смелость – робость). Высокие оценки свидетельствуют о смелости, решительности, пренебрежении опасностями, тяге к риску, легкости общения. Люди с низкими оценками робки, нерешительны, застенчивы, они избегают риска, опасностей.

**Фактор I** (мягкосердечность – жестокость). Высокие оценки имеют люди мягкие утонченные, с художественным восприятием мира. Низкие оценки являются показателем практичности, реалистичности, суровости.

**Фактор L** (подозрительность – доверчивость). Высокие оценки говорят о недоверчивости, завистливости, раздражительности, склонности к соперничеству. Низкие оценки свидетельствуют о бескорыстности, уступчивости, терпимости, отсутствии чувства собственной исключительности.

**Фактор M** (мечтательность – практичность). Люди с высокими оценками характеризуются поглощенностью своими идеями, интересом к абстрактным проблемам, восторженностью и легкой от-

влекательностью. Низкие оценки говорят о спокойствии, твердости и надежности в практических вопросах.

**Фактор N** (дипломатичность – наивность). Высокие оценки определяют людей изысканных, обладающих точным умом, эмоционально сдержанных и проницательных. Низкие оценки говорят о прямоте, бестактности, неконкретности ума и эмоциональной раскрепощенности.

**Фактор O** (склонность к чувству вины – самоуверенность). Люди с высокими оценками беспокояны, озабочены, подавлены, склонны к самоупрекам, не верят в свои силы. Низкие оценки характерны для лиц веселых, жизнерадостных, уверенных в своих успехах, бесстрашных.

**Фактор Q<sub>1</sub>** (радикальность – консервативность). Высокие оценки свидетельствуют о стремлении к новому, нетрадиционному, к неустоявшимся взглядам и терпимости к противоречиям. Низкие оценки – признак приверженности к устоявшимся мнениям, предпочтения проверенных способов решения проблем, отсутствия критики к авторитету.

**Фактор Q<sub>2</sub>** (самостоятельность – зависимость от группы). Высокие оценки говорят о независимости, самостоятельности, наличии собственной жизненной позиции. Низкие оценки показывают несамостоятельность, зависимость, ориентированность на групповое мнение, стремление к поддержке и одобрению со стороны окружающих.

**Фактор Q<sub>3</sub>** (высокий самоконтроль поведения – импульсивность). Люди с высокими оценками умеют хорошо контролировать свои эмоции и поведение, действуют планомерно и упорядоченно. Низкие оценки свидетельствуют о слабой воле, плохом самоконтроле, неумении организовывать свою деятельность.

**Фактор Q<sub>4</sub>** (напряженность – расслабленность). Высокие оценки указывают на напряженность, возбуждение, беспокойство, особенно при возникновении трудностей и препятствий. Люди с низкими оценками невозмутимы, довольны любым положением дел, не стремятся к переменам.

Исходя из описания, делаем вывод о том, что методика выбрана корректно: содержание факторов пригодно для оценки внутренних факторов адаптации (табл. 9.1).

#### **Методика «Изучение мотивов учебной деятельности студентов»**

Методика изучения мотивов учебной деятельности студентов разработана А.А.Реаном и В.А.Якуниным [77]. Испытуемым предлагается оценить по семибалльной шкале каждый из 16 представленных мотивов учебной деятельности. На основе произведенной оценки определяются 6 основных мотивов учебной деятельности:

- Приобретение знаний
- Овладение профессией
- Получение диплома
- Учебный процесс
- Одобрение окружающих
- Стипендия

В результате делается вывод о доминировании тех или иных мотивов, которые побуждают испытуемых учиться.

Исходя из описания, делаем вывод о том, что методика выбрана корректно: содержание шкал пригодно для оценки учебной и профессиональной мотивации студентов (табл. 9.1).

#### **Анализ документации**

- Для того чтобы оценить уровень интеллектуального развития студентов первокурсников, был проведён анализ результатов ЕГЭ по математике.
- Для анализа школьной подготовки первокурсников был проведён анализ результатов ЕГЭ по математике и физике.
- Для того чтобы оценить состояние здоровья студентов, был произведён анализ документации (медицинская справка 086). Все школьники на основании медицинского заключения распределяются на две группы: основную и специальную. Основным критерием для включения в ту или иную медицинскую группу являются уровень здоровья и функциональное состояние организма:

- 1 – студенты со специальной медицинской группой;
- 2 – студенты с основной медицинской группой;

### **Анализ учебной самостоятельности методом анкетного опроса**

Для определения уровня учебной самостоятельности методом анкетного опроса нами была разработана анкета, вопросы которой были сформулированы на основе следующих критериев сформированности учебной самостоятельности:

- Объем и качество знаний - характеризует потенциальные возможности личности студента в реализации самостоятельной учебной деятельности (показателем этого критерия может служить уровень успеваемости студента, то есть его экзаменационная оценка)
- Стабильность учения - характеризует студента с точки зрения качества организации познавательной и самостоятельной работы.
- Достаточность знаний – характеризует умение студента самостоятельно находить источники учебной информации и извлекать из них необходимые сведения, анализировать и оценивать факты, явления, выявлять причинно-следственные связи, обосновывать собственные суждения.
- Сформированность ценностного отношения к самостоятельной учебной деятельности – характеризует наличие у студента потребности в самостоятельной учебной деятельности, понимание важности и значимости этой деятельности в профессиональном становлении, умение адекватно оценивать собственную деятельность
- Активность в самостоятельной учебной деятельности как критерий характеризует созидательный и творческий характер этой деятельности, позволяющий не только пользоваться полученными знаниями, но и совершенствовать их.

Таким образом, подобранный набор методик позволяет оценить адаптационные возможности студентов, а также внутренние факторы успешности процесса социально-психологической адаптации.



### 9.2.4 Методы статистического анализа данных

Для проверки положения гипотезы о том, что студенты, успешно адаптировавшиеся к обучению в вузе, обладают сходным набором индивидуальных характеристик, имеющих различную выраженность в зависимости от курса обучения, нам необходимо из общей выборки студентов выделить группу (кластер) испытуемых, обладающих высокими показателями по измеренным психологическим характеристикам успешной адаптации: хорошая успеваемость, моральная нормативность, коммуникативные способности, нервно-психическая устойчивость. Для этой цели предполагается использовать **кластерный анализ**. Для определения «набора индивидуальных характеристик», которыми отличаются кластер адаптировавшихся студентов от остальных сокурсников предполагается использовать **U-критерий Манна-Уитни**. Для изучения структуры взаимосвязи этих характеристик предполагается использовать **коэффициент корреляции Спирмена**.

### 9.2.5 Этапы экспериментальной работы

Для достижения цели исследования и проверки экспериментальной гипотезы были определены следующие этапы экспериментальной работы:

1. Подобрать экспериментальную группу из студентов 1, 2 и 3 курса физико-энергетического факультета в количестве 100-120 человек.
2. Провести психодиагностическое обследование студентов с использованием следующих методик:
  - опросника «Адаптивность» (МЛО-АМ) (А.Г. Маклаков, С.В. Чермянин),
  - методики «Изучение мотивов учебной деятельности студентов» (А.А. Реан, В.А. Якунин);
  - опросника 16-ФЛО Р. Кеттелла.
3. Провести изучение уровня учебной самостоятельности студентов 1, 2 и 3 курса методом анкетного опроса.
4. Провести работу с документацией:

- проанализировать успеваемость студентов 1, 2 и 3 курса на основе рейтингового балла и результатов экзаменационных сессий;
- проанализировать результаты ЕГЭ по математике и физике студентов 1, 2 и 3 курса;
- проанализировать состояние здоровья студентов 1, 2 и 3 курса.

5. Методом кластерного анализа разделить общую выборку выборку на группы со схожими характеристиками успеваемости и адаптационных особенностей

6. Проанализировать значимость различий состояния здоровья, учебной самостоятельности, школьной подготовки, личностных особенностей, уровня развития интеллектуальных особенностей, учебной мотивации между категориями студентов 1, 2 и 3 курса с помощью критерия Манна-Уитни.

7. Изучить связи между показателями и факторами адаптации студентов 1, 2 и 3 курса с помощью коэффициента ранговой корреляции Спирмена.

8. Провести сравнительный анализ выраженности показателей и внутренних факторов адаптации у адаптировавшихся и неадаптировавшихся студентов 1, 2 и 3 курса на момент начала обучения

9. Провести сравнительный анализ показателей адаптации у студентов 2 и 3 курса, посещавших и не посещавших тренинги с помощью критерия Манна-Уитни

10. Сделать выводы и выработать научно-практические рекомендации по оптимизации процесса адаптации первокурсников к обучению в вузе.

### **9.2.6 Содержание экспериментальной работы**

Экспериментальное изучение внутренних факторов успешности процесса социально-психологической адаптации первокурсников к обучению в вузе проводилось в Обнинском институте атомной энергетики. В ходе работы были обследованы группы студентов первого, второго и третьего курсов, обучающихся по одинаковой программе. Выборка составила 101 человек, в возрасте 16-22 года.

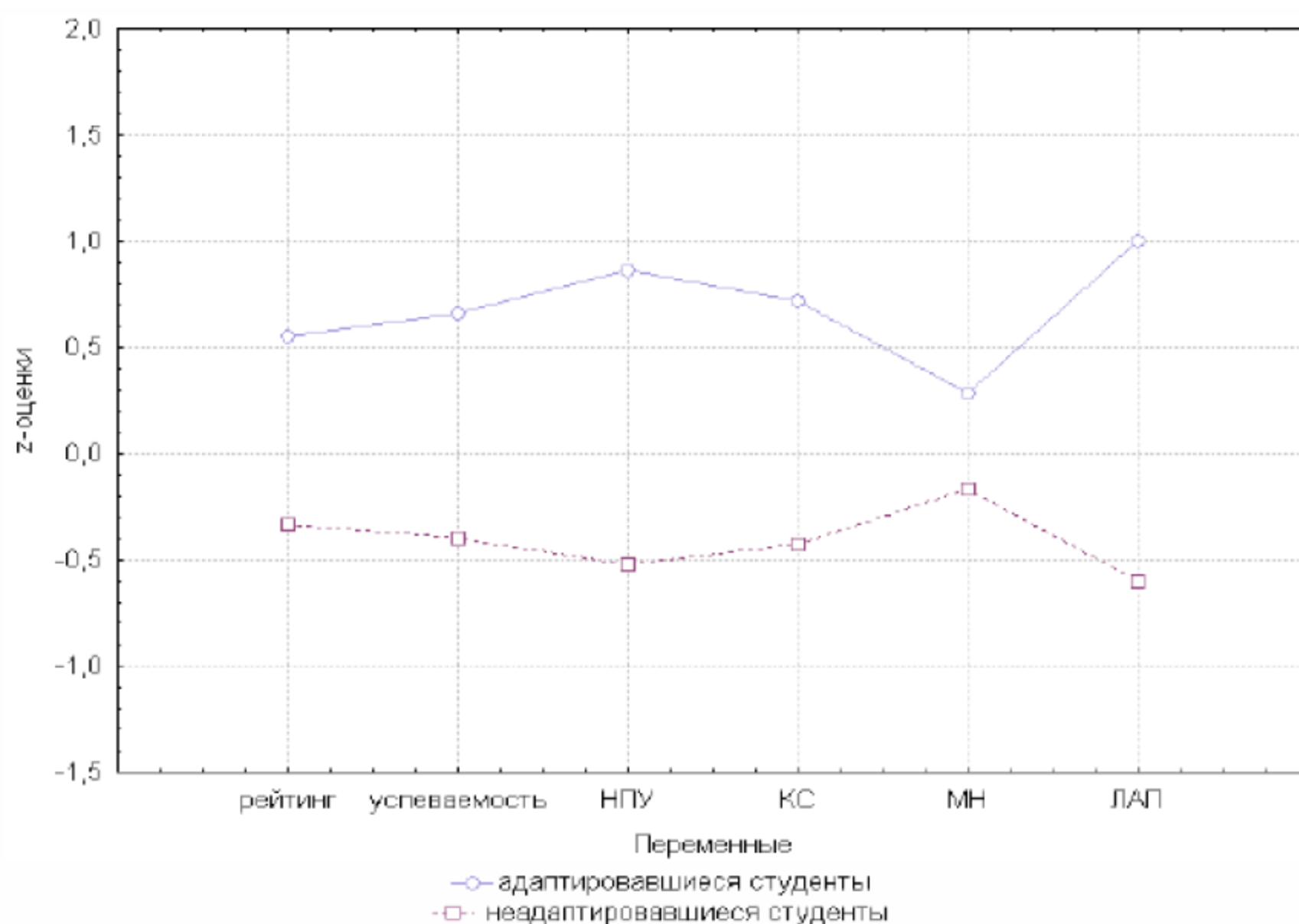
Для проверки положения гипотезы о том, что студенты, успешно адаптировавшиеся к обучению в вузе, обладают сходным набором индивидуальных характеристик, имеющих различную выраженность в зависимости от курса обучения, нами была отдельно проанализирована выраженность индивидуальных характеристик у студентов 1, 2 и 3 курса.

### **9.2.7 Статистический анализ результатов**

#### **Категории студентов в соответствии с уровнем критериев адаптации**

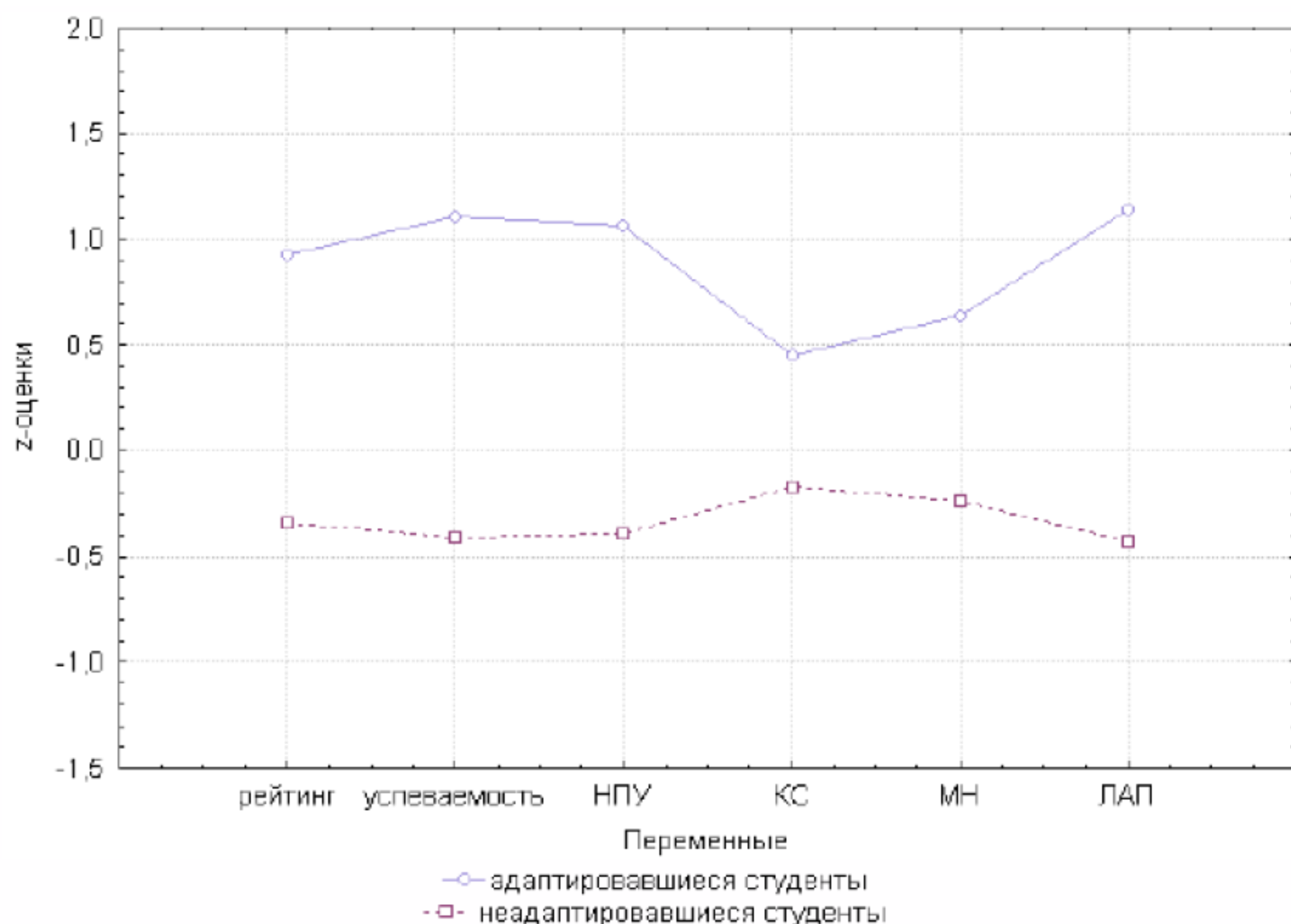
На первом этапе анализа данных методом кластерного анализа выборка студентов каждого курса была разделена на группы адаптировавшихся и неадаптировавшихся к обучению в вузе в соответствии с уровнем выраженности показателей адаптации (табл. 9.1).

В результате студенты 1 курса разделились на две группы: адаптировавшиеся и неадаптировавшиеся студенты (рис.9.1). Адаптировавшиеся студенты имеют высокую успеваемость и высокий рейтинговый балл, высокие оценки по шкалам НПУ (нервно-психическая устойчивость), КС (коммуникативные способности), МН (моральная нормативность) и ЛАП (личностный адаптивный потенциал). У неадаптировавшихся студентов успеваемость и рейтинговый балл по сравнению с адаптировавшимися достаточно низкий, показатели НПУ, КС, МН и ЛАП также находятся на низком уровне. С помощью критерия Манна-Уитни были выявлены статистически значимые различия между кластерами по следующим критериям адаптации: рейтинг ( $U=59$ ,  $p=0,02$ ), успеваемость ( $U=64$ ,  $p=0,03$ ), НПУ ( $U=24,5$ ,  $p=0,01$ ), КС ( $U=44$ ,  $p=0,01$ ) и ЛАП ( $U=12,5$ ,  $p=0,01$ ).



**Рис. 9.1 Показатели адаптации первокурсников, адаптировавшихся и неадаптировавшихся к обучению (данные стандартизированы)**

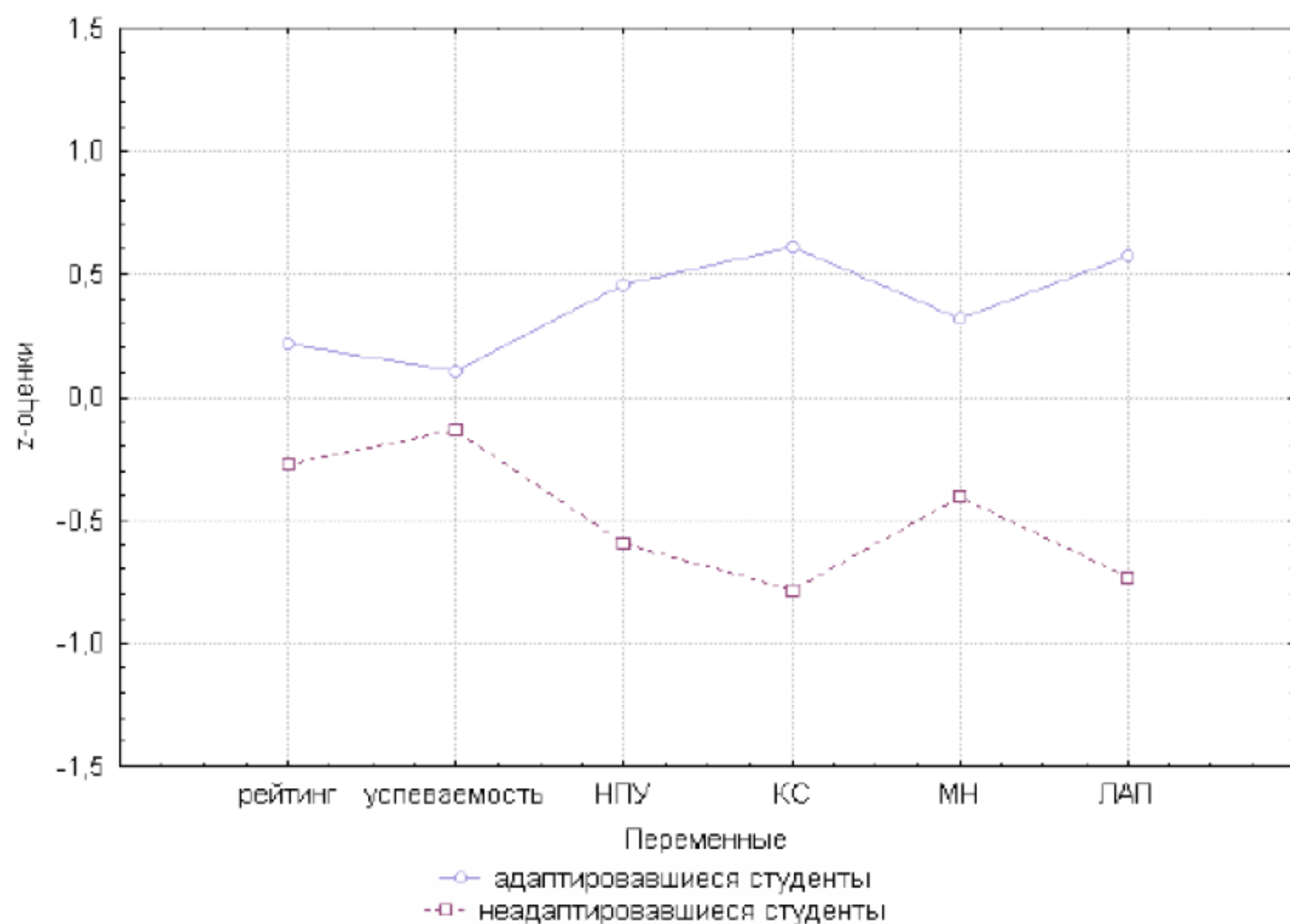
Студенты 2 курса также разделились на две группы (рис.9.2). Адаптировавшиеся студенты имеют высокую успеваемость и высокий рейтинговый балл, высокие оценки по шкалам НПУ (нервно-психическая устойчивость), КС (коммуникативные способности), МН (моральная нормативность) и ЛАП (личностный адаптивный потенциал). У неадаптировавшихся студентов успеваемость и рейтинговый балл достаточно низкий, показатели НПУ, КС, МН и ЛАП также находятся на низком уровне. Были выявлены статистически значимые различия между кластерами по следующим критериям адаптации: рейтинг ( $U=19$ ,  $p=0,01$ ), успеваемость ( $U=30$ ,  $p=0,01$ ), НПУ ( $U=19$ ,  $p=0,01$ ), МН ( $U=67$ ,  $p=0,02$ ) и ЛАП ( $U=17$ ,  $p=0,01$ ).



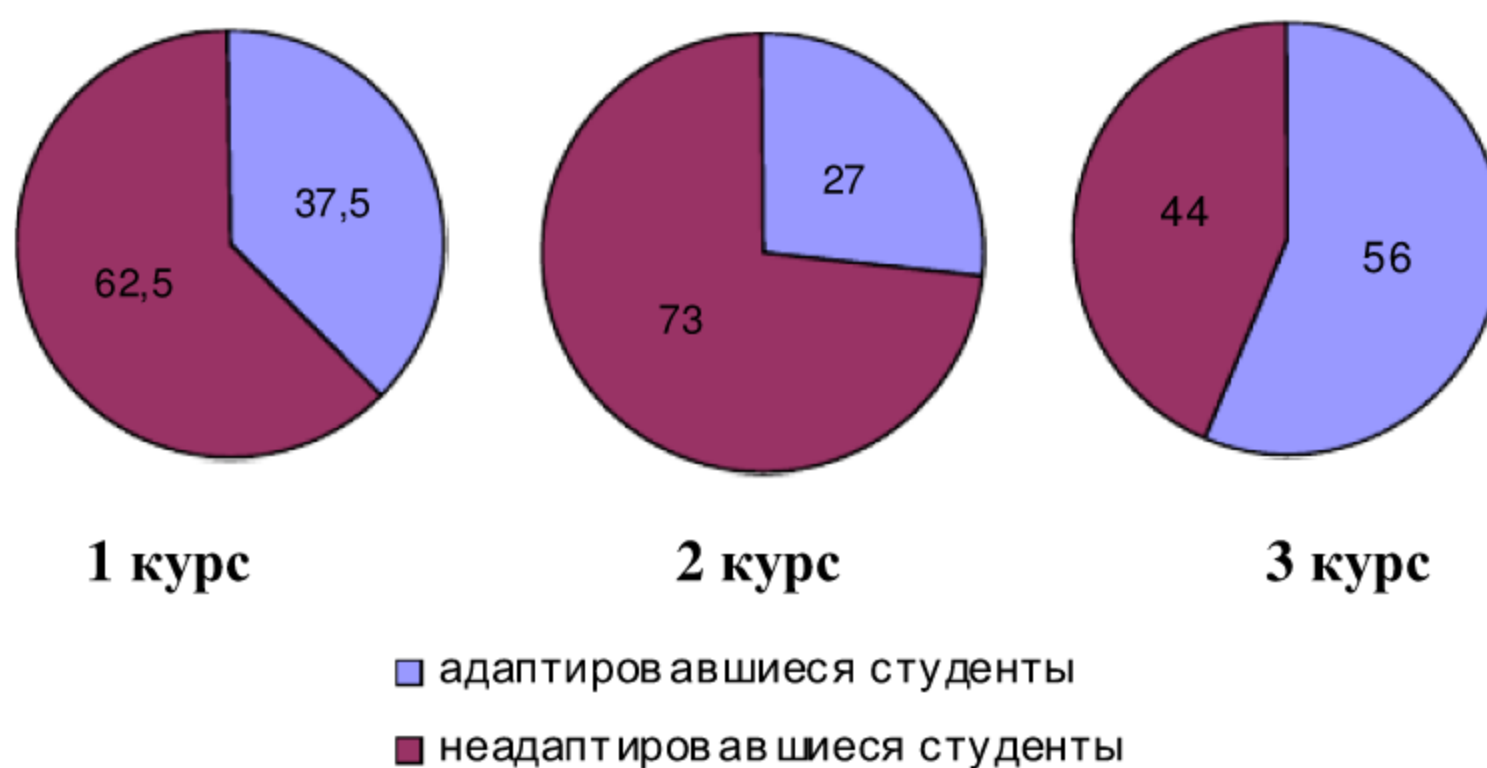
**Рис.9.2 Показатели адаптации студентов 2 курса, адаптировавшихся и неадаптировавшихся к обучению (данные стандартизированы)**

Студенты 3 курса также разделились на две группы (рис.9.3). Адаптировавшиеся студенты имеют высокую успеваемость и высокий рейтинговый балл, высокие оценки по шкалам НПУ (нервно-психическая устойчивость), КС (коммуникативные способности), МН (моральная нормативность) и ЛАП (личностный адаптивный потенциал). У неадаптировавшихся студентов успеваемость и рейтинговый балл по сравнению с адаптировавшимися достаточно низкий, показатели НПУ, КС, МН и ЛАП также находятся на низком уровне. Статистическая значимость различий между кластерами установлена по следующим критериям адаптации: НПУ ( $U=52,2$ ,  $p=0,01$ ), КС ( $U=27$ ,  $p=0,01$ ) и ЛАП ( $U=22$ ,  $p=0,01$ ).

На рис.9.4 представлены доли адаптировавшихся и неадаптировавшихся студентов в выборке первокурсников, второкурсников и третьекурсников.



**Рис.9.3 Показатели адаптации студентов 3 курса, адаптировавшихся и неадаптировавшихся к обучению (данные стандартизированы)**



**Рис. 9.4. Процентные доли адаптировавшихся и неадаптировавшихся студентов выборках своего курса**



По результатам анализа данных, представленных на рис. 9.4, сделаны следующие выводы:

- 37,5% первокурсников успешно адаптировались к обучению в вузе, обладают высокими психологическими показателями адаптации; 62,5% первокурсников не адаптировались к обучению, они обладают низкими психологическими показателями адаптации.
- 27% студентов 2 курса адаптировались к обучению, обладают высокими педагогическими и психологическими показателями адаптации; 73% студентов не адаптировались к обучению, и обладают низкими педагогическими и психологическими показателями адаптации. Результаты свидетельствуют о том, что доля неадаптировавшихся студентов 2 курса больше, чем доля неадаптировавшихся первокурсников, но она не достигла статистической значимости (что доказано с помощью углового преобразования Фишера-  $\phi^*$ ). Это, возможно, вызвано сложностью дисциплин, преподаваемых на втором курсе. В жизни второкурсников интенсивно включены все формы обучения и воспитания. Второй курс - период самой напряженной учебной деятельности студентов. К тому же в конце семестра напряжение может быть вызвано тем, что у студентов в это время проходят контрольные работы. Помимо того, близятся сроки сдачи различных заданий, ведь сессия не за горами. Также на 2-3 курсах студенты проходят кризис юношеского возраста (18-19 лет). Суть его состоит в том, что меняются отношения с миром: реализация своих юношеских идеалов требует иного, не созерцательного, а действенного отношения к жизни. Не имея достаточного жизненного опыта и ресурсов, некоторые студенты оказываются в состоянии дезадаптации. К аналогичным выводам приходил В.В. Белехов, обследовавший студентов ФЭФ ИАТЭ в 1983-1993 годах [3].
- 56% студентов 3 курса адаптировались к обучению, обладают высокими педагогическими и психологическими показателями адаптации; 44% студентов не адаптировались к обучению, и

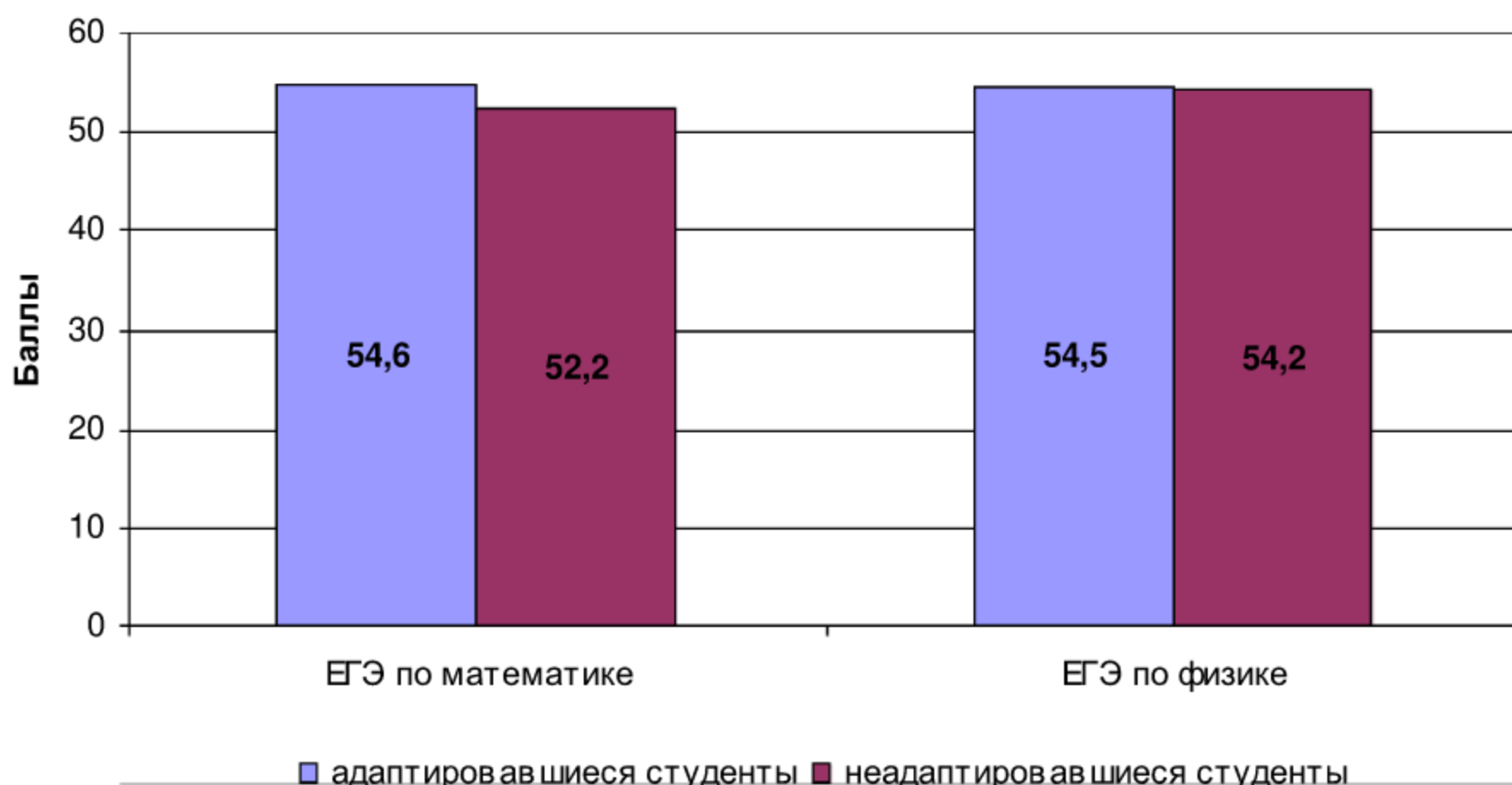
обладают низкими педагогическими и психологическими показателями адаптации. По педагогическим показателям адаптации студенты не различаются, у обеих групп студентов успеваемость достаточно высокая. В процессе обучения студенты овладевают способами учения и к третьему курсу могут концентрироваться непосредственно на изучаемом материале, а не на формах обучения. Результаты свидетельствуют о том, что доля неадаптировавшихся студентов 3 курса статистически значимо уменьшилась по сравнению со 2 курсом (что доказано с помощью углового преобразования Фишера –  $\phi^* = 2,49$ ,  $p = 0,005$ ). Однако большая часть студентов так и не смогла адаптироваться к обучению, в первую очередь, по психологическим показателям адаптации. Как уже отмечалось выше, на 2-3 курсах студенты проходят кризис юношеского возраста, что может привести к дезадаптации.

### **Сопоставление выраженности внутренних факторов успешности адаптации у студентов**

Анализ выраженности внутренних факторов адаптации первокурсников, в зависимости от принадлежности к кластерам успешности адаптации, проводился с помощью критерия Манна-Уитни. В результате было установлено, что

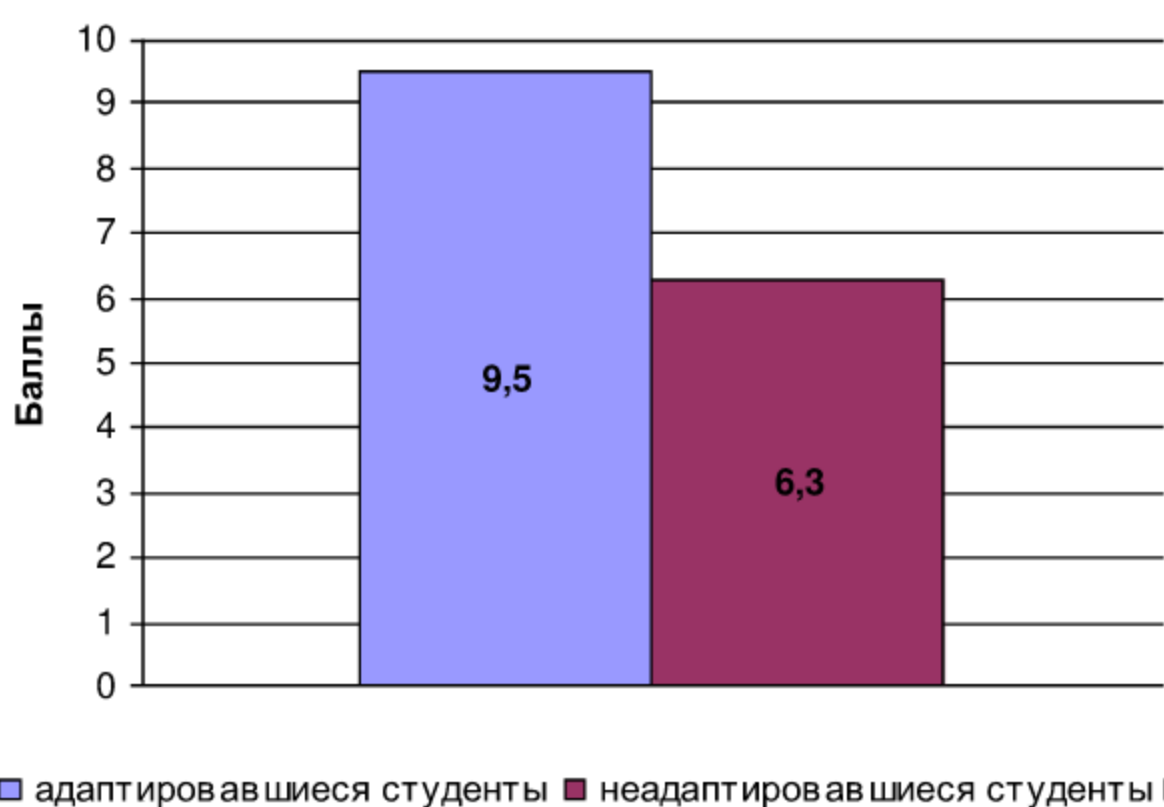
а) Состояние здоровья не является определяющим фактором успешности адаптации: статистически значимых различий в долях студентов специальной медицинской группы в выборках адаптировавшихся и неадаптировавшихся первокурсников не выявлено.

б) Статистически значимых различий между показателями школьной подготовки у адаптировавшихся и неадаптировавшихся студентов также не выявлено (рис.9.5). Анализ данных, представленных на диаграмме, показывает, что у адаптировавшихся студентов результаты ЕГЭ по математике и физике несколько выше, чем у неадаптировавшихся, но они не достигли статистической значимости.

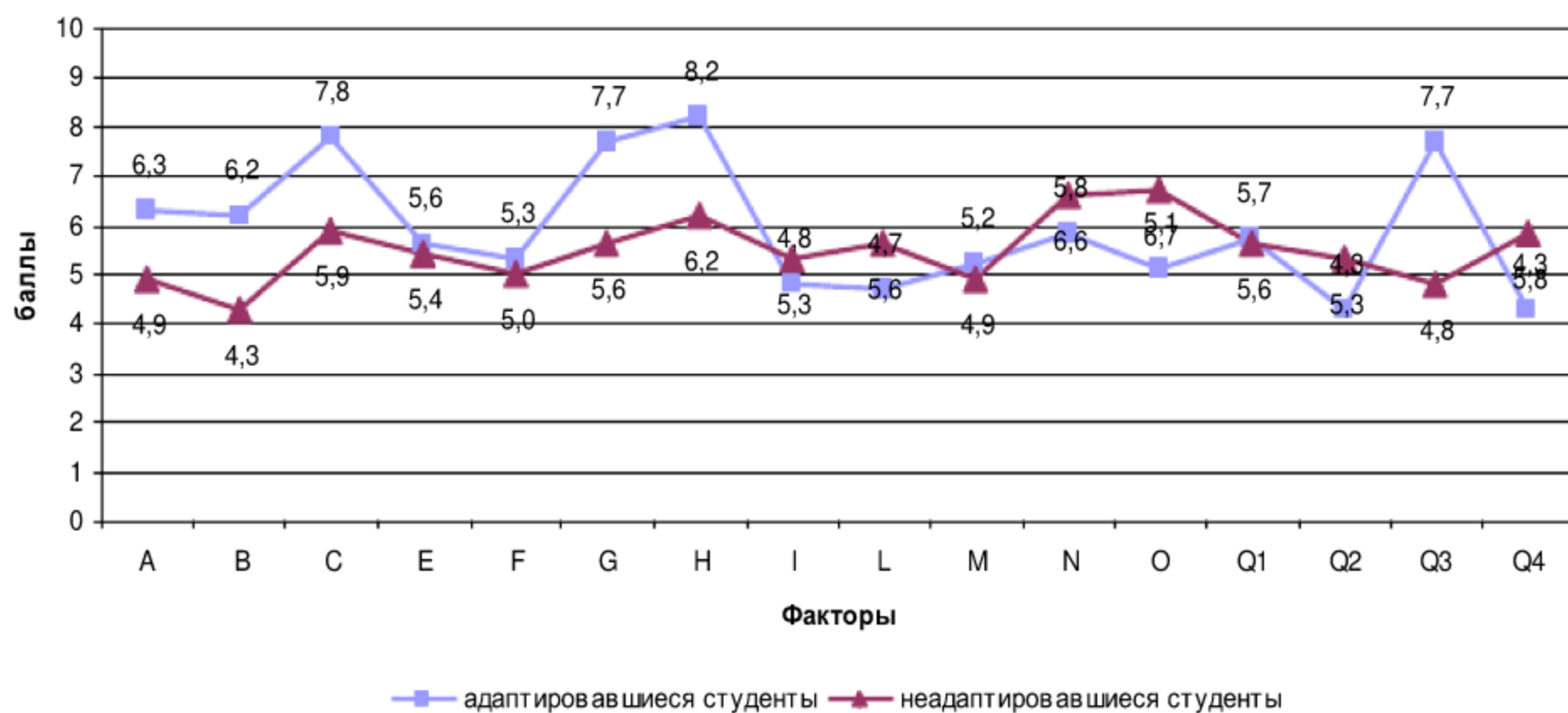


**Рис. 9.5 Уровень школьной подготовки у адаптировавшихся и неадаптировавшихся студентов 1 курса**

в) Уровень учебной самостоятельности адаптировавшихся первокурсников достоверно выше ( $U=31,5$ ,  $p=0,01$ ), чем у неадаптировавшихся (рис.9.6). Таким образом, адаптировавшиеся студенты умеют рационально распределять время, силы, планировать свою работу.



**Рис. 9.6. Уровень учебной самостоятельности у адаптировавшихся и неадаптировавшихся студентов 1 курса**



**Рис. 9.7 Личностные особенности адаптировавшихся и неадаптировавшихся студентов 1 курса (по 16 ФЛО Кеттелла)**

г) В результате сравнительного анализа показателей личностных особенностей адаптировавшихся и неадаптировавшихся студентов было выявлено следующее (рис.9.7):

- существуют статистически значимые различия по фактору В («высокий интеллект - низкий интеллект») ( $U=52,5$ ,  $p=0,01$ ). Фактор В (интеллектуальная активность) является косвенным показателем мотивации выполнения заданий теста. То есть адаптировавшиеся студенты обладают гибким аналитичным умом, сообразительностью и высокой обучаемостью. Неадаптировавшимся студентам свойственны конкретность мышления, плохая способность к обучению.
- существуют статистически значимые различия по фактору С («эмоциональная устойчивость – эмоциональная неустойчивость») ( $U=52,5$ ,  $p=0,01$ ). То есть адаптировавшимся студентам свойственны эмоциональная зрелость, уверенность в себе, работоспособность (кстати, это подтверждают и данные методики МЛЮ-АМ). Неадаптировавшиеся студенты менее эмоционально зрелые и работоспособные, чем адаптировавшиеся.

- существуют статистически значимые различия по фактору G («высокая совестливость – недобросовестность») ( $U=49$ ,  $p=0,02$ ). Адаптировавшиеся студенты дисциплинированы, добросовестны, во всем любят порядок. Неадаптировавшиеся студенты менее дисциплинированы, непостоянны.
- существуют статистически значимые различия по фактору H («смелость – робость») ( $U=39$ ,  $p=0,01$ ). Адаптировавшиеся студенты обладают хорошими коммуникативными навыками (кстати, это подтверждают и данные методики МЛЮ-АМ), уверены в себе; особенно сильно раскрываются эти качества в ситуации соперничества и риска. Неадаптировавшиеся студенты менее решительные и активные.
- существуют статистически значимые различия по фактору Q<sub>3</sub> («высокий самоконтроль поведения – импульсивность») ( $U=20,5$ ,  $p=0,01$ ). Адаптировавшиеся студенты лучше управляют собственными эмоциями. Различия по факторам говорят о стремлении к упорядоченности и осознанности ценностей и мотивов. Неадаптировавшиеся студенты обладают слабой волей и плохим самоконтролем.
- существуют статистически значимые различия по фактору Q<sub>4</sub> («напряженность – расслабленность») ( $U=59,5$ ,  $p=0,03$ ). Адаптировавшиеся студенты довольны положением дел, спокойны. Для неадаптировавшихся студентов характерны напряженность, беспокойство, особенно при возникновении трудностей и препятствия.

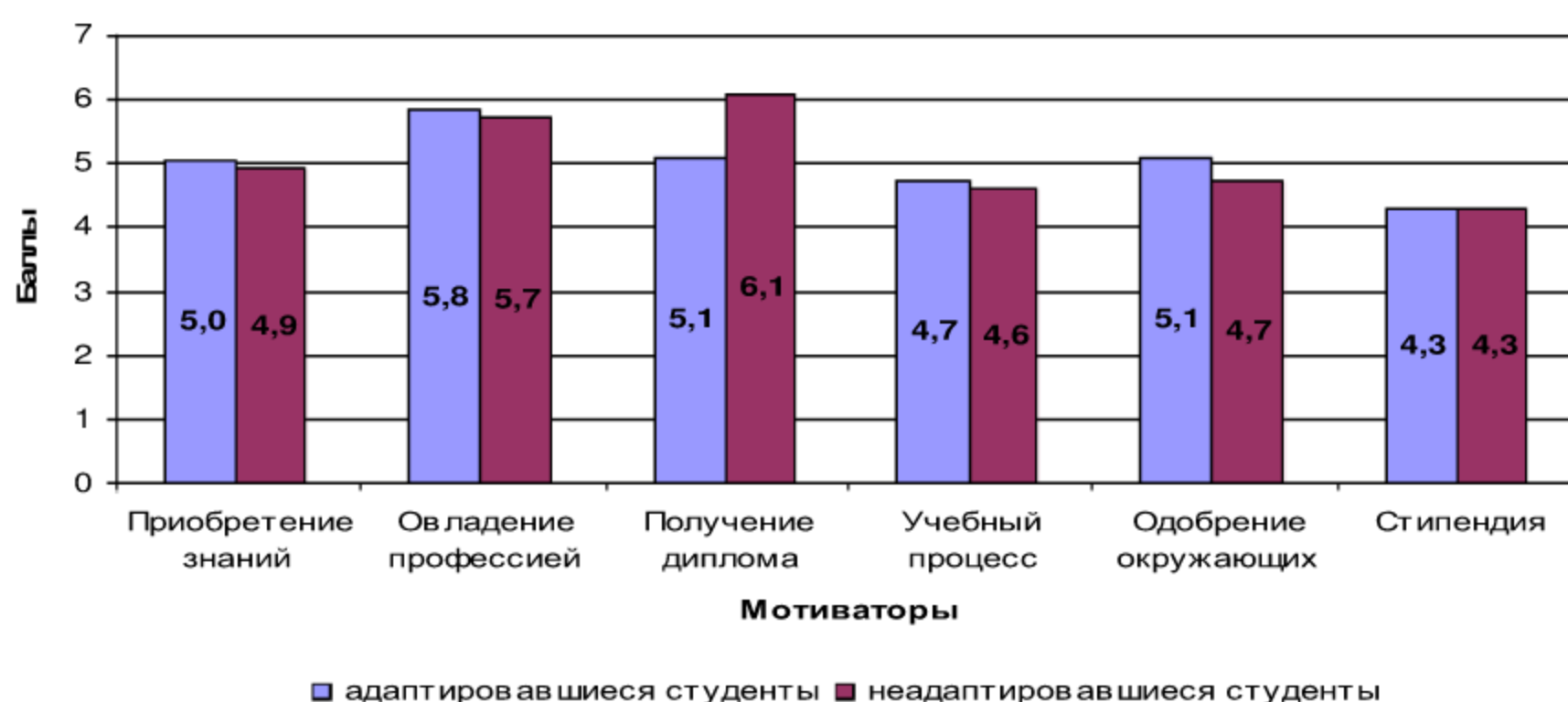
д) В результате сравнительного анализа интеллектуального развития студентов (в качестве показателей были выбраны результаты ЕГЭ по математике и фактор В 16-факторного личностного опросника Кеттелла) было установлено:

- отсутствие статистически значимых различий между результатами ЕГЭ по математике у адаптировавшихся и неадаптировавшихся студентов.
- существование статистически значимые различия по фактору В ( $U=52,5$ ,  $p=0,01$ ). Содержание этого фактора характеризует ин-



теллектуальную активность испытуемого. Поэтому можно утверждать, что адаптировавшиеся первокурсники демонстрируют более высокую интеллектуальную активность по сравнению с неадаптировавшимися.

е) Различия по показателям мотивации у адаптировавшихся и неадаптировавшихся студентов (рис. 9.8) не являются статистически значимыми.



**Рис.9.8 Показатели мотивации у адаптировавшихся и неадаптировавшихся студентов 1 курса**

Таким образом, успешно адаптировавшиеся первокурсники обладают следующим набором индивидуальных характеристик:

- высокий уровень учебной самостоятельности
- высокий уровень интеллектуальной активности
- личностные особенности: эмоциональная устойчивость, уверенность в себе, работоспособность, дисциплинированность, добросовестность, решительность, высокий самоконтроль эмоций и поведения.

Аналогичным образом был проведён сравнительный анализ выраженности внутренних факторов успешности процесса у студентов тех же специальностей 2 и 3 курса, в результате которого были получены следующие выводы:

Успешно адаптировавшиеся студенты 2 курса обладают следующим набором индивидуальных характеристик:



- высокий уровень учебной самостоятельности
- высокая внутренняя учебная мотивация (мотивы «учебный процесс» и «получение стипендии»)
- личностные особенности: дисциплинированность, добросовестность, высокий самоконтроль эмоций и поведения, терпимость и уживчивость.

Успешно адаптировавшиеся студенты 3 курса обладают следующим набором индивидуальных характеристик:

- личностные особенности: эмоциональная яркость и устойчивость, активность, дисциплинированность и добросовестность, смелость, дипломатичность, проницательность, расслабленность, высокий самоконтроль эмоций и поведения.

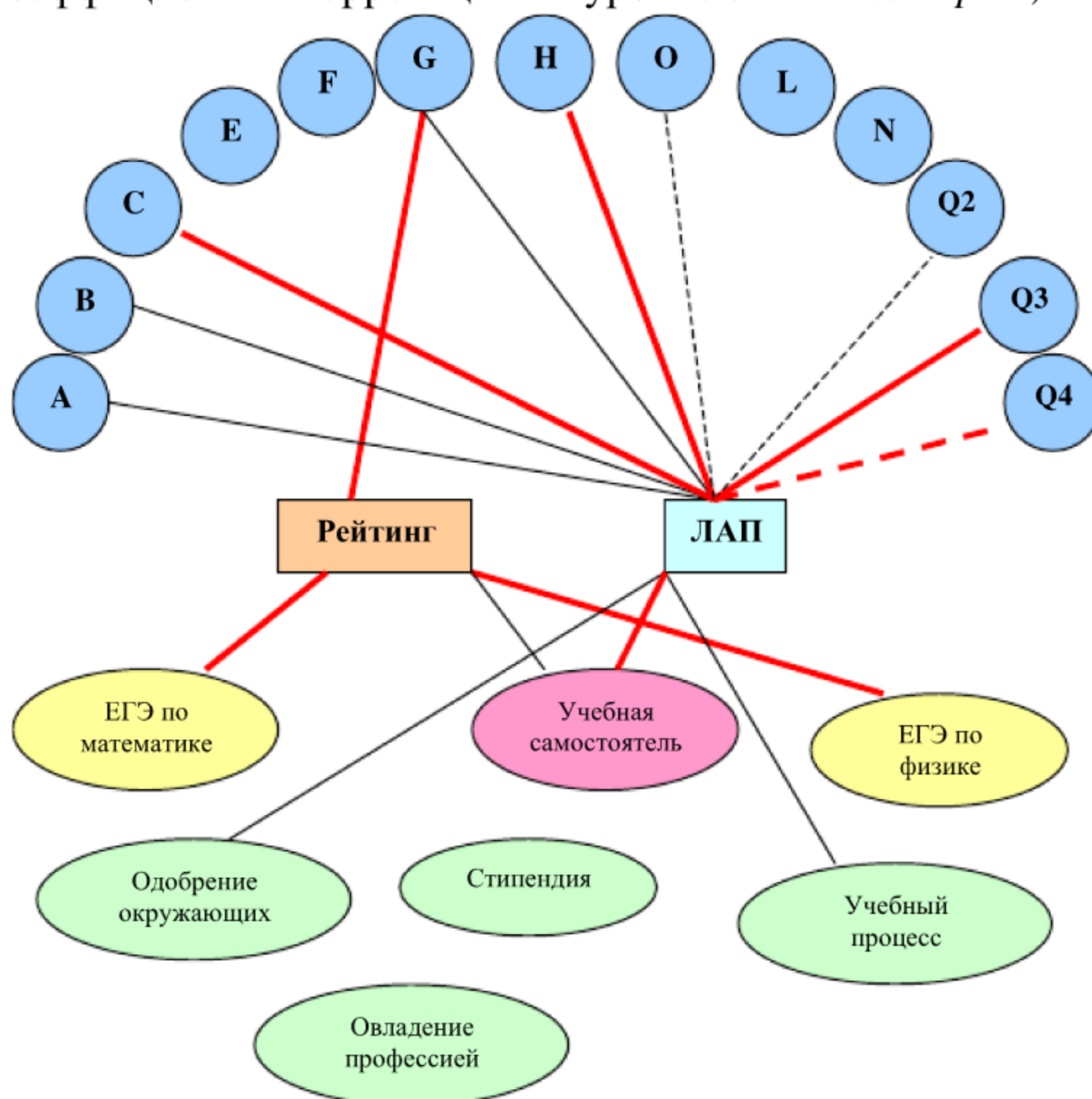
Таким образом, сравнительный анализ индивидуальных характеристик адаптировавшихся и неадаптировавшихся студентов 1,2 и 3 курса ФЭФ показал, что определяющими внутренними факторами успешности адаптации являются, личностные характеристики (в первую очередь, эмоциональная устойчивость и дисциплинированность), а также интеллектуальная активность и мотивация для студентов 1-2 курсов. Школьная подготовка и состояние здоровья не вошли в число внутренних факторов, определяющих успешность адаптации.

### **Взаимовязь показателей и внутренних факторов адаптации студентов**

Для изучения особенностей взаимосвязи показателей и внутренних факторов адаптации студентов были введены два интегративных показателя адаптации: педагогическим показателем адаптации явился рейтинговый балл, психологическим – показатель личностного адаптивного потенциала (ЛАП). Взаимосвязь этих интегративных показателей с внутренними факторами адаптации (табл. 9.2) изучалась с помощью коэффициента корреляции Спирмена.

На рис. 9.9 сплошными линиями показаны прямые корреляционные связи, пунктирными линиями – обратные связи. Более тонкие линии изображают корреляционные связи между показателями

на уровне значимости  $p < 0,05$ , а более толстые линии соответствуют коэффициентам корреляции на уровне значимости  $p < 0,01$ .



**Рис. 9.9. Корреляционные связи показателей адаптации с внутренними факторами адаптации у студентов 1 курса**

Как показывает анализ корреляционного графа (рис. 9.9), педагогический показатель адаптации «Рейтинг» имеет среднюю прямую корреляционную связь с факторами ЕГЭ по математике  $r=0,55$  и ЕГЭ по физике  $r=0,64$ , фактором G  $r=0,68$ ; умеренную прямую связь с фактором Учебная самостоятельность  $r=0,45$ .

Таким образом, педагогический показатель адаптации первокурсника прямо коррелирует:

- с уровнем школьной подготовки;
- с учебной самостоятельностью;
- с личностными особенностями: фактором добросовестности G.

Психологический показатель адаптации «ЛАП» имеет среднюю прямую связь с факторами С  $r=0,5$ , Н  $r=0,69$ , Q3  $r=0,59$  и Учебная самостоятельность и среднюю обратную связь с фактором Q4  $r=0,62$ ; умеренную прямую связь с факторами А  $r=0,42$ , В  $r=0,5$ , G  $r=0,38$ , мотивами «Учебный процесс»  $r=0,35$  и «Одобрение окружающих»  $r=0,34$  и умеренную обратную связь с факторами О  $r=-0,46$  и Q2  $r=-0,36$ .

Следовательно, психологический показатель адаптации первокурсников имеет прямую взаимосвязь

- с учебной самостоятельностью;
- с мотивами учебной деятельности: «учебный процесс» и «одобрение окружающих»;
- с личностными особенностями: факторами доброжелательности А, интеллектуальной активности В, эмоциональной стабильности С, добросовестности G, смелости Н, организованности Q3.

Кроме того, психологический показатель адаптации первокурсников имеет обратную взаимосвязь с факторами тревожности О, социабельности Q2 и фрустрируемости Q4.

Взаимосвязь показателей и внутренних факторов адаптации студентов 2 курса представлена на корреляционном графе (рис.9.10). Сплошными линиями на рисунке показаны прямые корреляционные связи, пунктирными линиями – обратные связи. Более тонкие линии изображают корреляционные связи между показателями на уровне значимости  $p<0,05$ , а более толстые линии соответствуют коэффициентам корреляции на уровне значимости  $p<0,01$ .

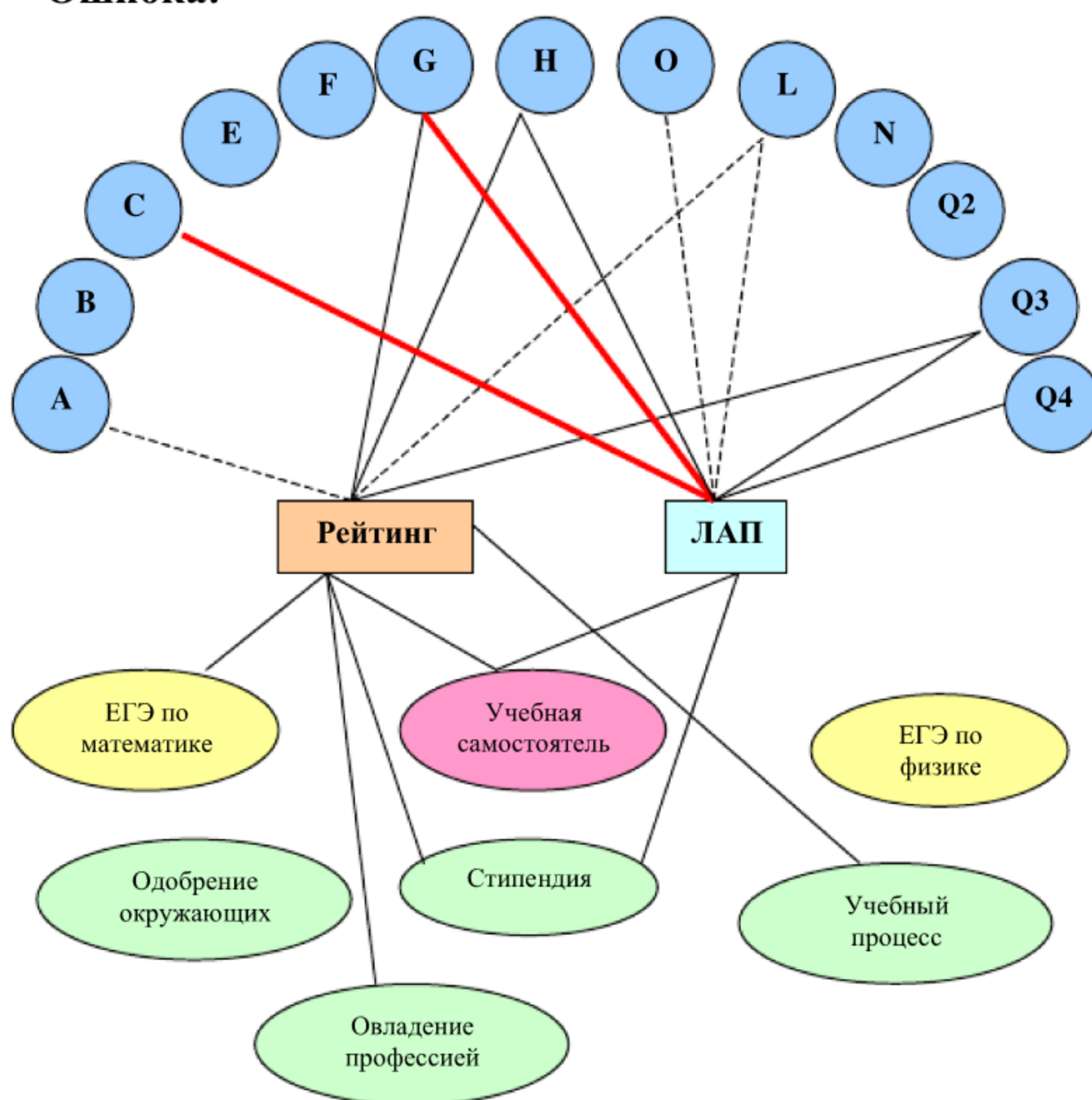
Как показывает анализ корреляционного графа (рис. 9.10), педагогический показатель адаптации «Рейтинг» имеет умеренную прямую корреляционную связь с результатом ЕГЭ по математике  $r=0,4$ , фактором G  $r=0,49$ , Н  $r=0,49$ , Q3  $r=0,39$ , мотивами «Овладение профессией»  $r=0,45$ , «Учебный процесс»  $r=0,38$  и «Стипендия»  $r=0,47$ , Учебная самостоятельность  $r=0,49$ ; умеренную обратную связь с факторами А  $r=-0,35$ , L  $r=-0,33$ .

Таким образом, педагогический показатель адаптации второкурсников имеет прямую взаимосвязь

- с уровнем школьной подготовки (ЕГЭ по математике);
- с учебной самостоятельностью;
- с мотивами учебной деятельности: «овладение профессией», «учебный процесс» и «получение стипендии»;
- с личностными особенностями: факторами добросовестности G, смелости H, организованности Q3.

Педагогический показатель адаптации второкурсников имеет обратную корреляционную взаимосвязь с личностными особенностями: факторами доброжелательности A и подозрительности L.

**Ошибка!**



**Рис. 9.10. Корреляционные связи показателей адаптации с внутренними факторами адаптации у студентов 2 курса**

Психологический показатель адаптации «ЛАП» имеет среднюю прямую связь с факторами С  $r=0,52$ , G  $r=0,51$ ; умеренную прямую связь с факторами Н  $r=0,44$ , Q3  $r=0,49$ , мотивом «Стипендия»  $r=0,41$  и Учебная самостоятельность  $r=0,36$  и умеренную обратную связь с факторами L  $r=0,-452$  O  $r=-0,34$ .

Таким образом, личностный адаптивный потенциал студента второкурсника прямо коррелирует:

- с учебной самостоятельностью;
- с мотивом учебной деятельности: «получение стипендии»;
- с личностными особенностями: факторами эмоциональной стабильности С, добросовестности G, смелости Н, организованности Q3.

Личностный адаптивный потенциал студента второкурсника обратно коррелирует с личностными особенностями: факторами подозрительности L, подавленности O и фрустрируемости Q4.

Взаимовязь показателей и внутренних факторов адаптации студентов 3 курса представлена на корреляционном графе (рис.9.11). Сплошными линиями на рисунке показаны прямые корреляционные связи, пунктирными линиями – обратные связи. Более тонкие линии изображают корреляционные связи между показателями на уровне значимости  $p<0,05$ , а более толстые линии соответствуют коэффициентам корреляции на уровне значимости  $p<0,01$ .

Как показывает анализ корреляционного графа (рис. 9.11), педагогический показатель адаптации «Рейтинг» имеет среднюю прямую корреляционную связь с факторами С  $r=0,55$ , G  $r=0,57$ , Н  $r=0,58$ , Q3  $r=0,57$ ; умеренную прямую связь с факторами Е  $r=0,41$ , F  $r=0,41$ , Учебная самостоятельность  $r=0,38$ , мотивами «Одобрение окружающих»  $r=0,41$  и «Стипендия»  $r=0,39$ ; умеренную обратную связь с факторами O  $r=-0,42$  и Q4  $r=-0,46$ .

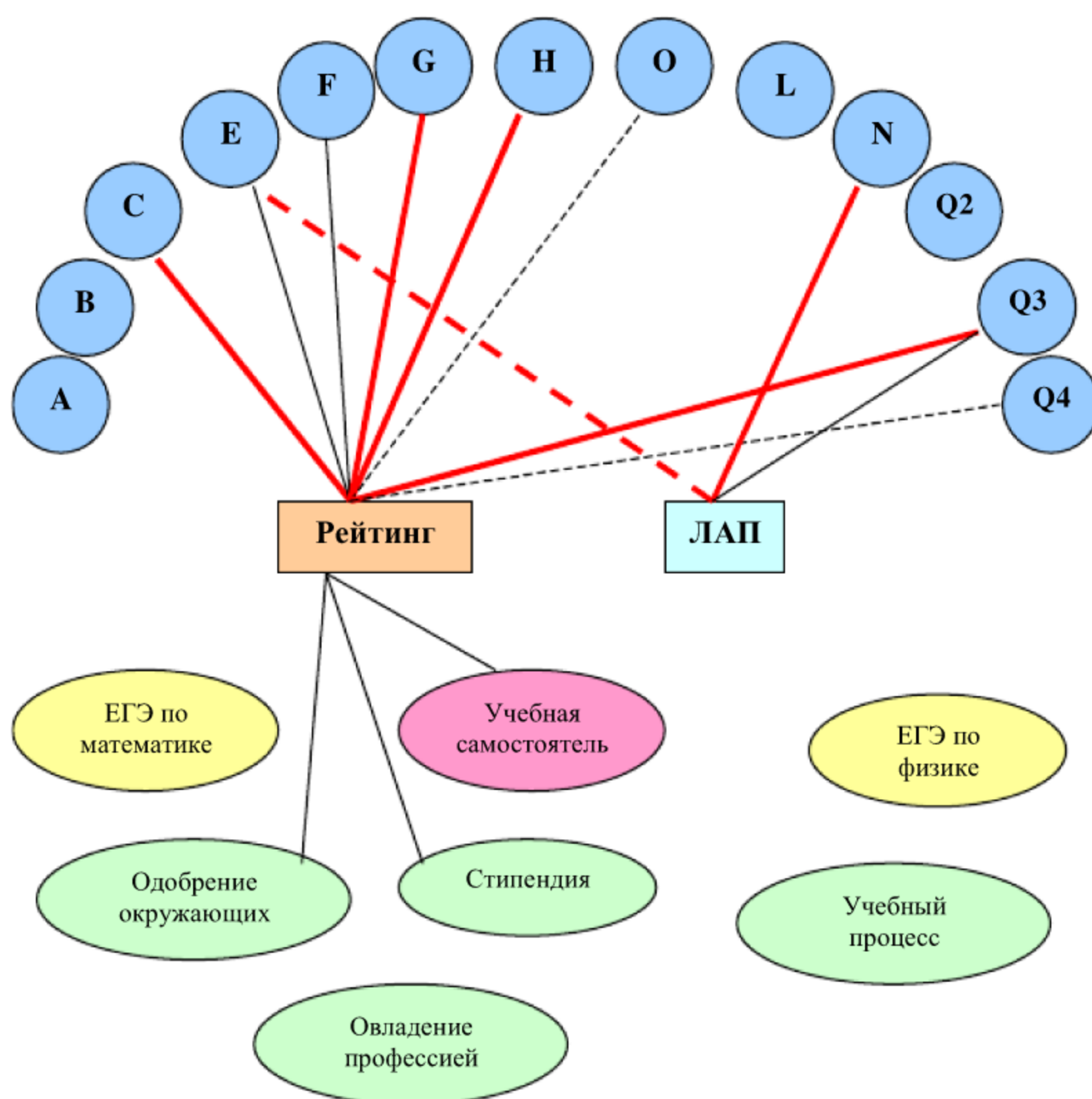
Таким образом, педагогический показатель адаптации третьекурсников имеет прямую взаимосвязь:

- с учебной самостоятельностью;



- с мотивами учебной деятельности: «одобрение окружающих» и «получение стипендии»;
- с личностными особенностями: факторами эмоциональной стабильности С, независимости Е, беспечности F, добросовестности G, смелости Н, организованности Q3.

Помимо этого, педагогический показатель адаптации третьекурсников имеет прямую взаимосвязь обратно коррелирует с фактором тревожности О и фрустрируемости Q4.



**Рис. 9.11. Корреляционные связи показателей адаптации с внутренними факторами адаптации у студентов3 курса**

Психологический показатель адаптации «ЛАП» имеет среднюю прямую связь с фактором N  $r=0,55$  и среднюю обратную связь с



фактором Е  $r=-0,56$ ; умеренную прямую связь с фактором Q3  $r=0,41$ . Следовательно, психологический показатель адаптации студентов 3 курса прямо коррелирует с личностными особенностями: фактором проницательности N и организованности Q3 и обратно коррелирует с фактором независимости Е.

Таким образом, если у студентов 1 курса показатель адаптации «Рейтинг» прямо коррелирует, в первую очередь, с результатами ЕГЭ по физике и математике и ответственностью студентов (фактор G), то на последующих курсах рейтинг связан в большей степени с неинтеллектуальными компонентами индивидуальности студентов, в первую очередь, с личностными особенностями: эмоциональной устойчивостью, социальной смелостью, ответственностью, организованностью, и на всех курсах – с учебной самостоятельностью. Психологический показатель адаптации студентов 1-2 курса связан как с личностными особенностями студентов, так и с показателями мотивации и учебной самостоятельности. У студентов 3 курса он связан исключительно с личностными особенностями. Следовательно, к 3 курсу адаптивные возможности студентов становятся частью их личностной структуры.

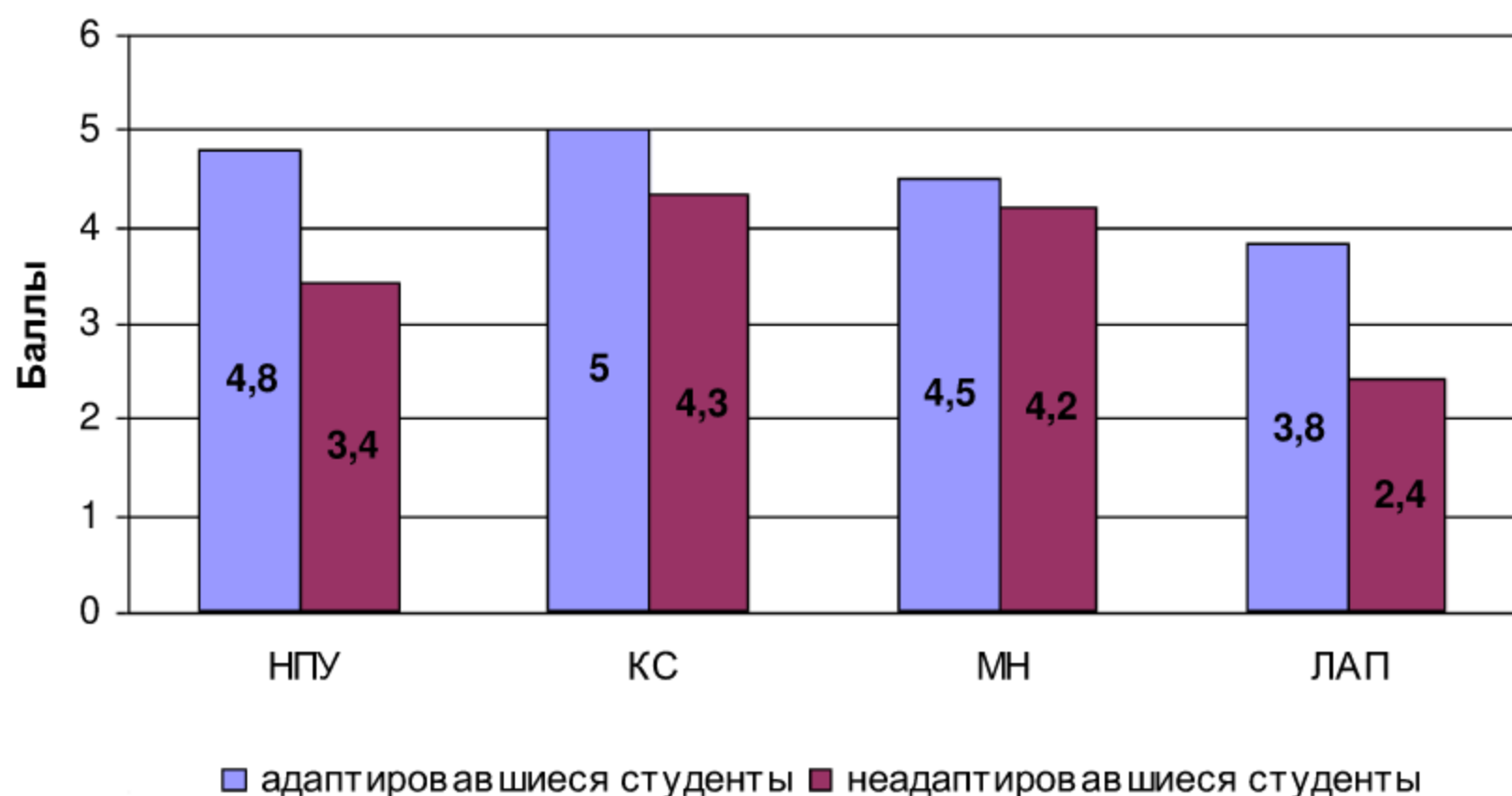
### **Сформированность внутренних факторов адаптации у студентов на момент начала обучения**

Для проверки положения гипотезы о том, что часть набора индивидуальных характеристик студентов, успешно адаптировавшихся к обучению, сформирована уже к моменту поступления в вуз, мы сравнили выраженность показателей адаптации и набора индивидуальных характеристик у адаптировавшихся и неадаптировавшихся студентов 1, 2 и 3 курса на момент начала обучения.

Сопоставление показателей адаптации у адаптировавшихся и неадаптировавшихся студентов 1-3 курсов с помощью критерия Манна-Уитни привело к следующим выводам.

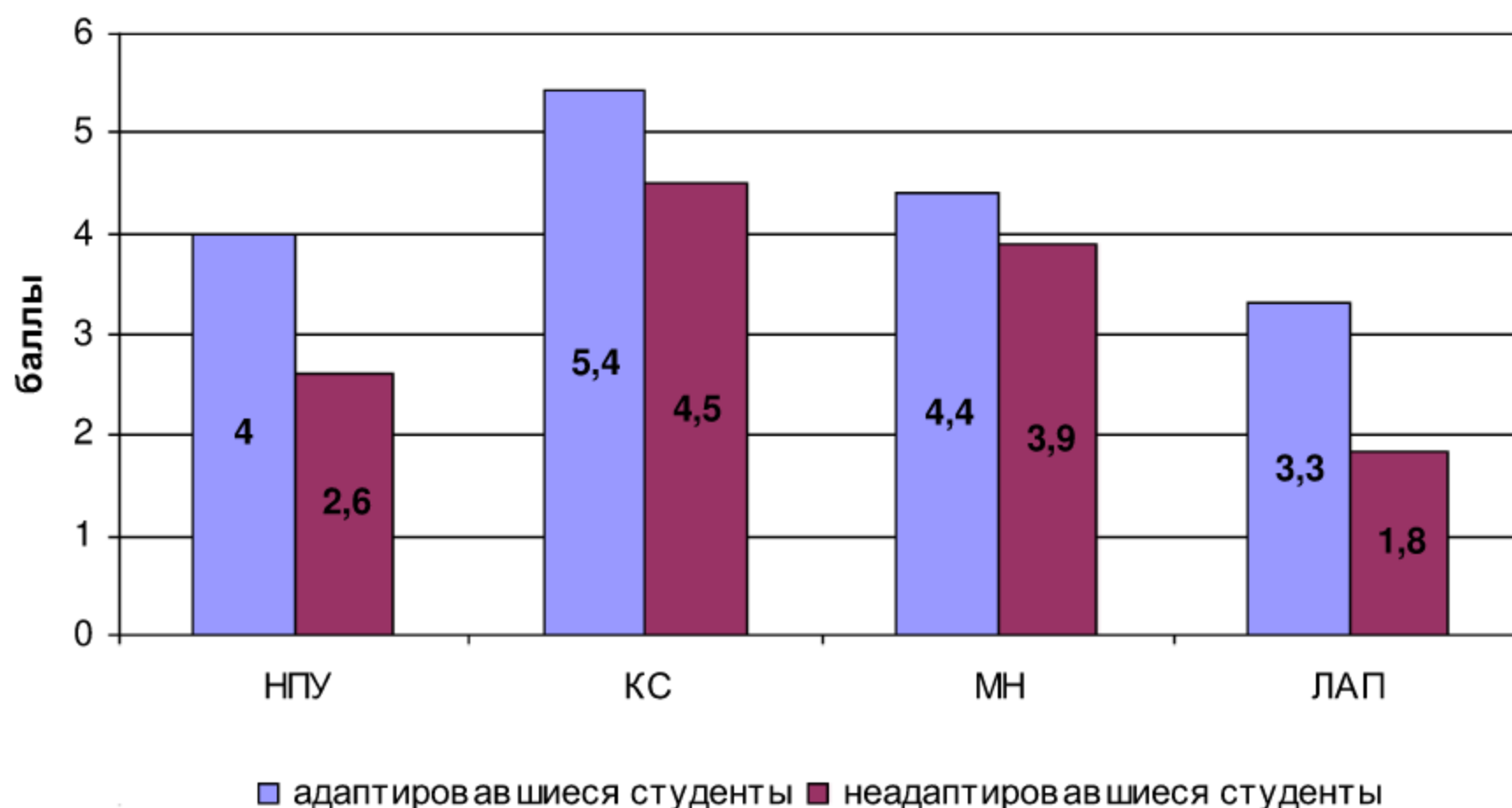
1. Существуют статистически значимые различия между адаптировавшимися и неадаптировавшимися первокурсниками по шкале нервно-психической устойчивости ( $p=0,01$ ) и интегративному показателю ЛАП ( $p=0,01$ ), которые у адаптировавшихся студентов

1 курса на момент начала обучения были выше, чем у неадаптировавшихся (рис. 9.12).



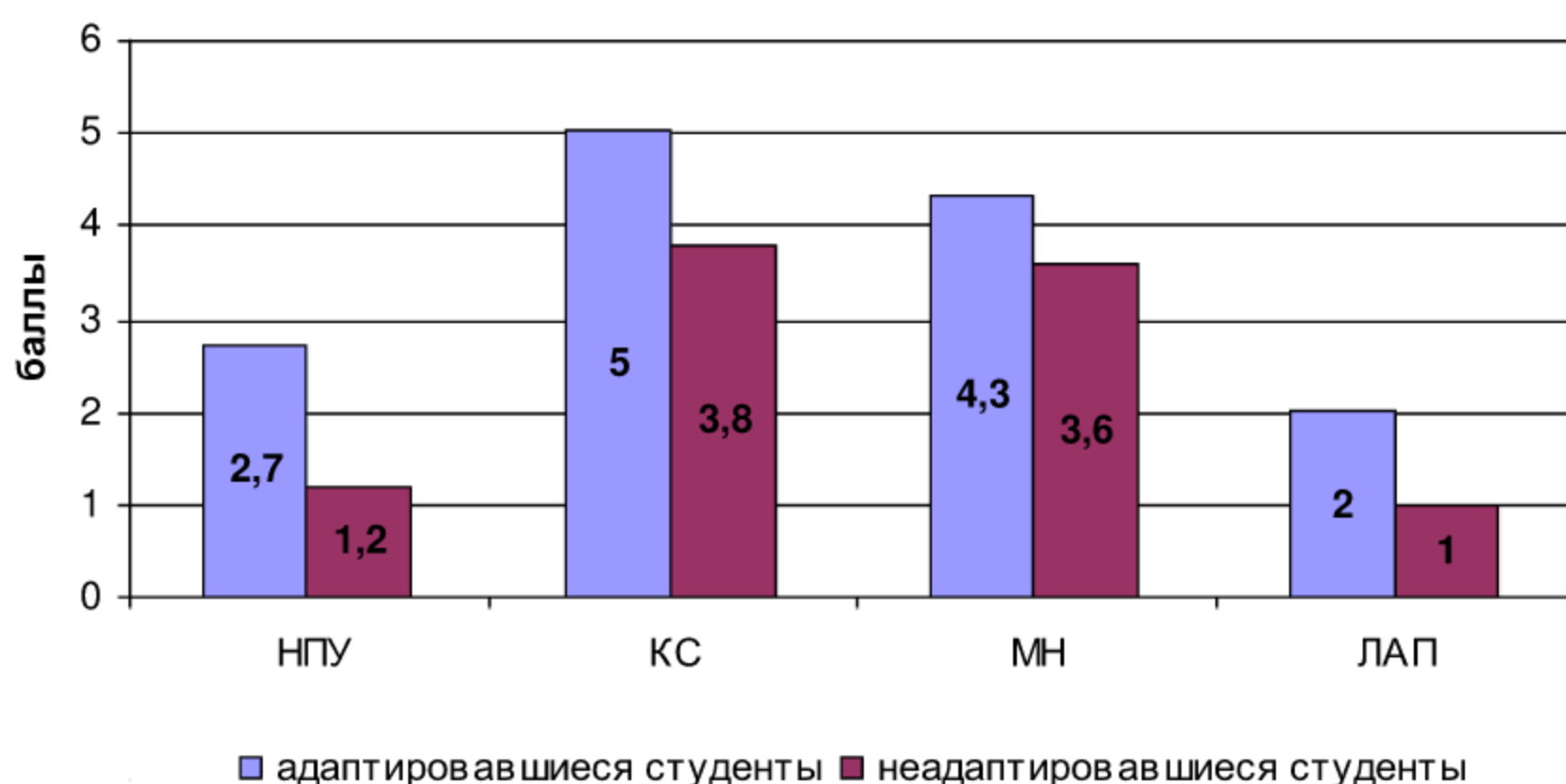
**Рис. 9.12 Показатели адаптации у студентов 1 курса на момент начала обучения.**

2. Существуют статистически значимые различия по шкалам нервно-психической устойчивости ( $p=0,03$ ), коммуникативных способностей ( $p=0,03$ ) и интегративному показателю ЛАП ( $p=0,05$ ), которые у адаптировавшихся студентов 2 курса на момент начала обучения были выше, чем у неадаптировавшихся (рис. 9.13).



**Рис. 9.13 Показатели адаптации у студентов 2 курса на момент начала обучения.**

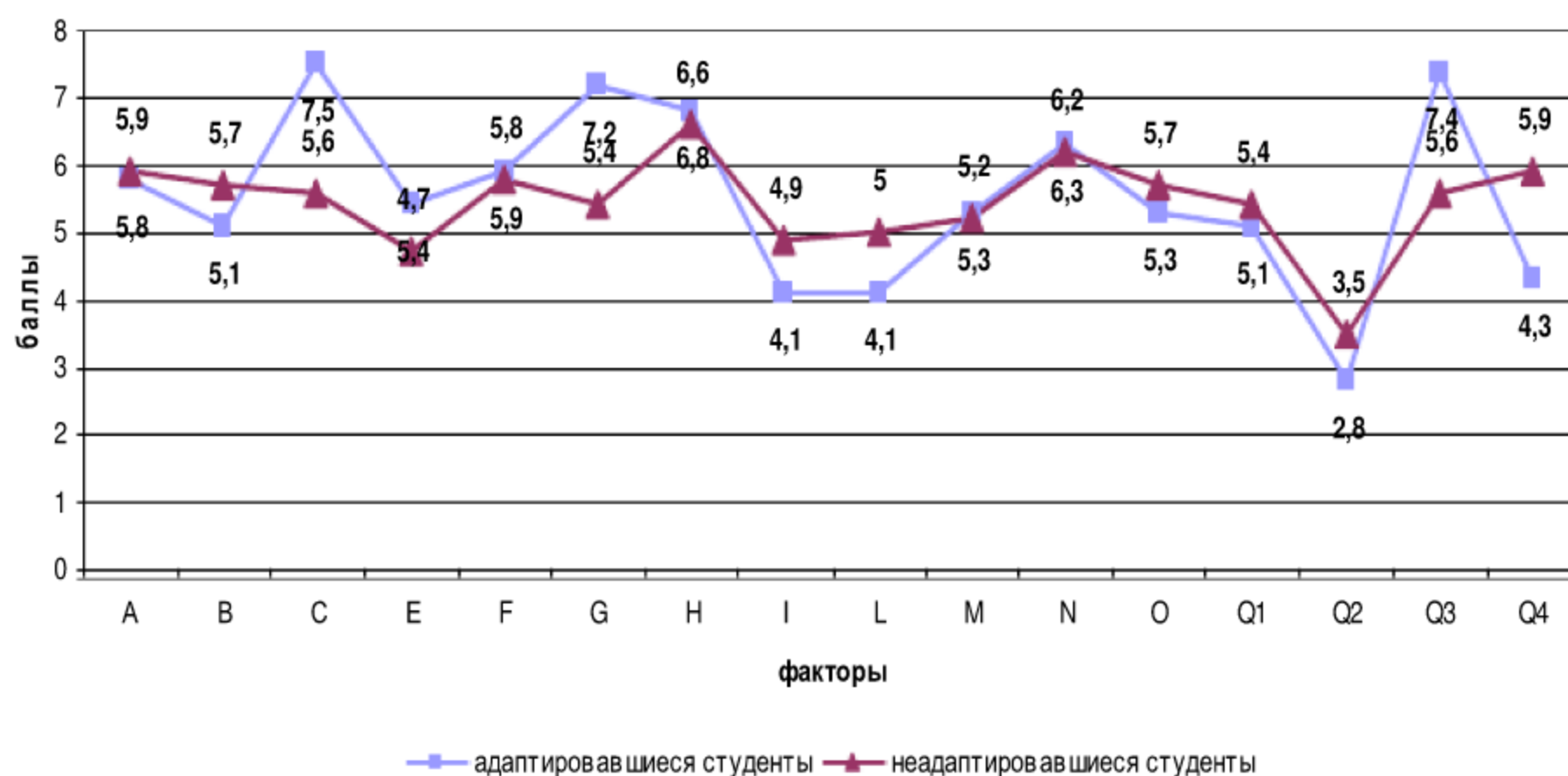
3. Существуют статистически значимые различия по шкалам нервно-психической устойчивости ( $p=0,04$ ), коммуникативных способностей ( $p=0,028$ ) и личностного адаптивного потенциала ( $p=0,02$ ). Показатели по этим шкалам у адаптировавшихся студентов 3 курса на момент начала обучения были выше, чем у их неадаптировавшихся однокурсников (рис. 9.14).



**Рис. 9.14 Показатели адаптации студентов 3 курса на момент начала обучения.**

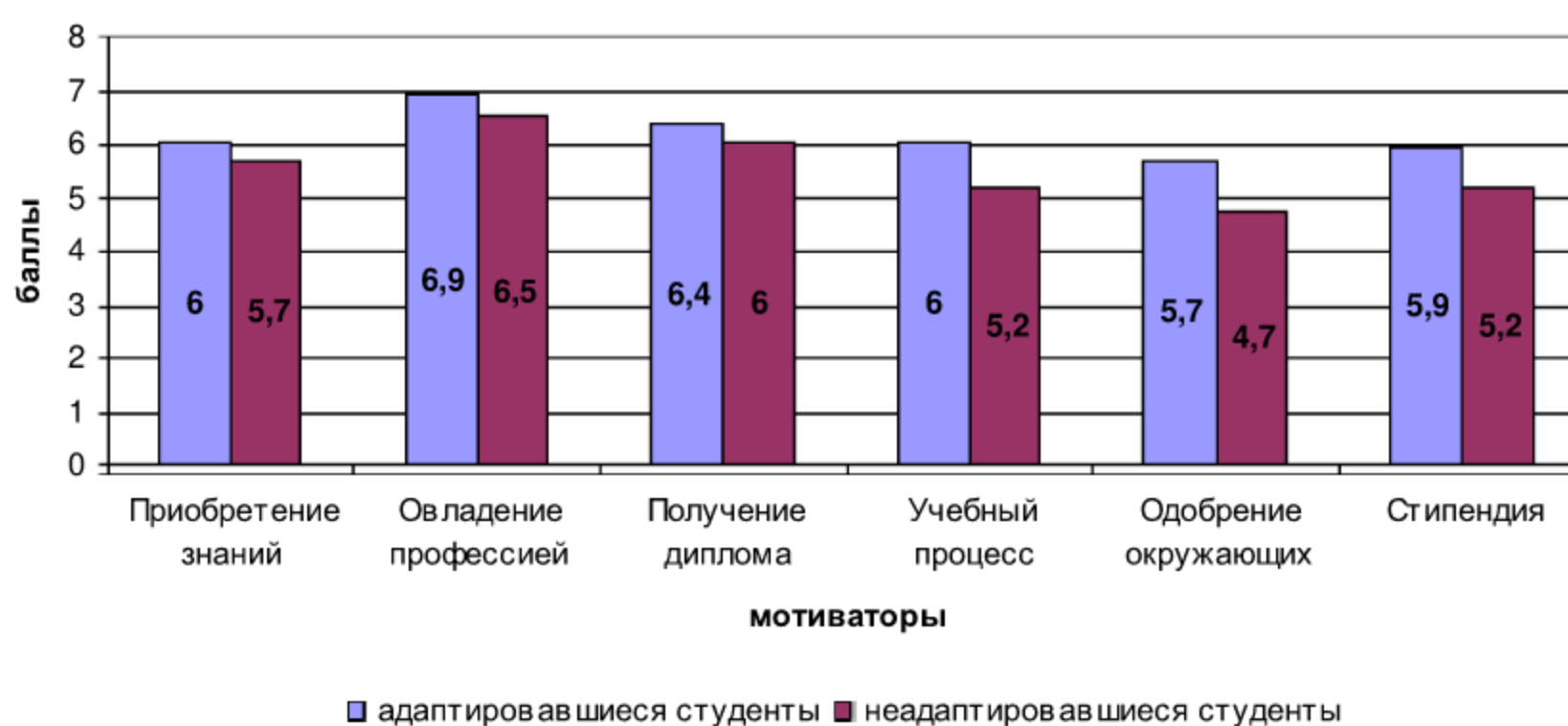
Сравнительный анализ личностных характеристик (на момент начала обучения в вузе) адаптировавшихся и не адаптировавшихся студентов 2 курса в показал (рис.9.15), что существуют достоверные различия по факторам С «эмоциональная устойчивость – эмоциональная неустойчивость» ( $p=0,01$ ), G «высокая совестливость – недобросовестность» ( $p=0,01$ ) и Q<sub>3</sub> «высокий самоконтроль поведения – импульсивность» ( $p=0,01$ ). То есть адаптировавшимся второкурсникам ещё на момент начала обучения в вузе были свойственны эмоциональная зрелость, уверенность в себе, работоспособность, дисциплинированность, добросовестность, высокий самоконтроль. Неадаптировавшиеся студенты на момент поступления

были менее эмоционально зрелыми и работоспособными, отличались непостоянством, неорганизованностью и недисциплинированностью.



**Рис. 9.15 Личностные характеристики (по 16 ФЛЮ Кеттелла) студентов 2 курса на момент начала обучения**

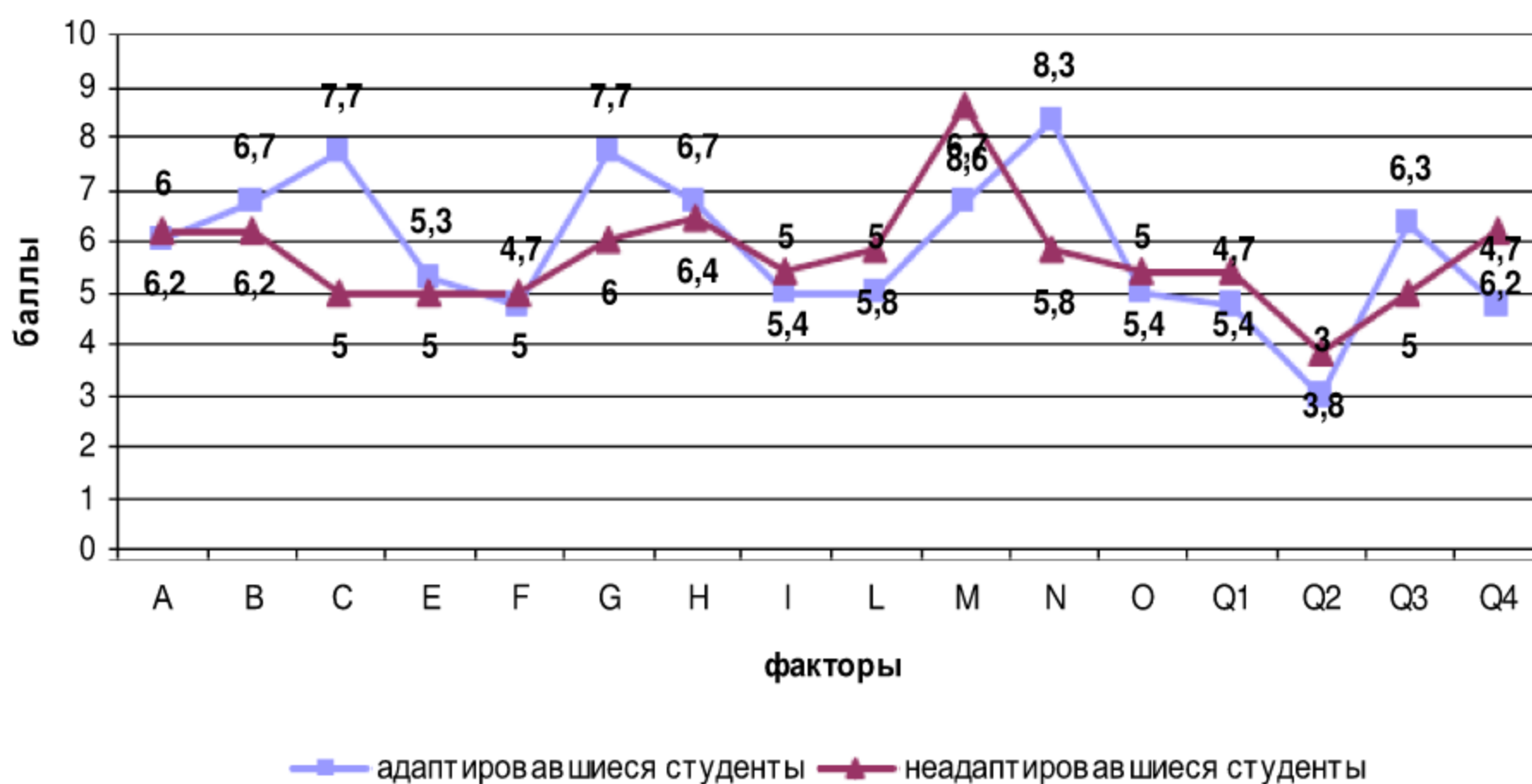
По показателям мотивации выявились следующие результаты (рис.9.16):



**Рис. 9.16 Показатели мотивации студентов 2 курса на момент начала обучения.**

Адаптировавшиеся студенты имеют статистически значимо более высокие показатели по мотиву учебной деятельности «учебный процесс» ( $p=0,03$ ). По остальным мотивам учебной деятельности у неадаптировавшихся студентов показатели немного ниже, но эта разница статистически не значима.

Сравнительный анализ личностных характеристик (на момент начала обучения в вузе) адаптировавшихся и не адаптировавшихся студентов 3 курса в показал (рис.9.16), что существуют статистически значимые различия по фактору С («эмоциональная устойчивость – эмоциональная неустойчивость») ( $p=0,02$ ) и G («высокая совестливость – недобросовестность») ( $p=0,04$ ). То есть адаптировавшимся студентам 3 курса ещё на момент начала обучения в вузе были свойственны эмоциональная зрелость, уверенность в себе, работоспособность, дисциплинированность, добросовестность, в отличие от их неадаптировавшихся однокурсников.



**Рис. 9.17 Личностные характеристики (по 16 ФЛЮ Кеттелла) студентов 3 курса на момент начала обучения**

Таким образом, в результате проведённого исследования доказано, что студенты, успешно адаптировавшиеся к обучению в вузе,



обладали следующим набором индивидуальных характеристик к моменту поступления в вуз: высокой стрессоустойчивостью, адекватной самооценкой, коммуникативными навыками, эмоциональной устойчивостью, дисциплинированностью и высоким самоконтролем поведения.

Учёт в образовательном процессе выявленных по результатам исследования внутренних факторов успешности адаптации студентов к обучению в вузе позволит оптимизировать процесс адаптации студентов и сократить его сроки. Необходимо проводить дополнительную психологическую подготовку кураторов и социально-активных студентов для проведения занятий по психологической поддержке, включающих в себя тренинговые занятия по развитию таких личностных особенностей как коммуникативные способности, решительность, дисциплинированность и добросовестность, контроль над эмоциями, развитие адекватной самооценки и снижение уровня тревожности.

В результате проведенного исследования решены все поставленные задачи, достигнута цель исследования, выдвинутая гипотеза получила экспериментальное подтверждение. Доказано, что студенты, успешно адаптировавшиеся к обучению в вузе, обладают сходным набором индивидуальных характеристик, имеющих различную выраженность в зависимости от курса обучения, при этом часть из этих характеристик сформирована у них уже к моменту поступления в вуз.



## Литература

1. Абрамова В. Н., Белехов В. В., Бельская Е. Г. и др. Психологические методы в работе с кадрами на АЭС. М., 1990.
2. Аванесов В. С. Тесты в социологическом исследовании. М., 1982.
3. Адаптация первокурсников к обучению в вузе: Сборник научных трудов /Под ред. Е. В. Леоновой. Обнинск: ИАТЭ, 2008.
4. Анастаси А., Урбина С. Психологическое тестирование. СПб., 2001.
5. Бодалев А.А., Столин В.В. Общая психодиагностика. СПб., 2000.
6. Болотова А.К., Макарова И.В. Прикладная психология: Учебник для вузов. М., 2001.
7. Боровиков В. Statistica: искусство анализа данных на компьютере. Для профессионалов. СПб., 2001.
8. Боровиков В. Программа Statistica для студентов и инженеров. М., 2001.
9. Бурлакова Н.С., Олешкевич В.И. Проективные методы: теория, практика применения к исследованию личности ребёнка. М., 2001.
10. Бурлачук Л.Ф. Психодиагностика: Учебник для вузов. СПб., 2003.
11. Бурлачук Л.Ф., Морозов С.М. Словарь-справочник по психодиагностике. СПб., 1999.
12. Выготский Л.С. Диагностика развития и педологическая клиника трудного детства. //Собр. соч. М.: Педагогика, 1983. Т.5. С. 257-321.
13. Габараев Б. А., Черепнин Ю. С. и др. Проблемы подготовки специалистов для атомной энергетики //Безопасность АЭС и

- подготовка кадров. X Международная конференция: Тезисы докладов. Часть 1. Обнинск, 2007. С.59-60.
14. Гальперин П. Я. Методы обучения и умственного развития ребенка. М.: Изд-во МГУ, 1985.
  15. Ганзен В. А. Системные описания в психологии. Л.: ЛГУ, 1984.
  16. Гласс Дж., Стэнли Дж. Статистические методы в педагогике и психологии. М., 1976.
  17. Гуревич К. М. Проблемы дифференциальной психологии. – М.; Воронеж, 1998.
  18. Гуревич К. М. Профессиональная пригодность и основные свойства нервной системы. М., 1970.
  19. Гуревич К. М. Что такое психологическая диагностика. М., 1985.
  20. Дружинин В. Н. Психология общих способностей. С-Пб., 2000.
  21. Дружинин В. Н. Экспериментальная психология. С-Пб., 2000.
  22. Ермолаев О.Ю. Математическая статистика для психологов. М., 2002.
  23. Зеер Э. Ф. Психология профессионального развития. М.: Академия, 2007.
  24. Зимняя И. А. Ключевые компетенции новая парадигма результата образования //Высшее образование сегодня. 2003. № 5. С.35-41.
  25. Ильин Е. П. Успешность деятельности, компенсации и компенсаторные отношения //Вопросы психологии. 1983. № 5. С.95-98.
  26. Канаев И.И. Фрэнсис Гальтон. Л., «Наука». 1972.
  27. Карандашев В.Н. Квалификационные работы по психологии: реферативные, курсовые и дипломные: Учеб.-метод. пособие. М.: Смысл, 2007.
  28. Кимбл Г. Как правильно пользоваться статистикой. М., 1982.
  29. Клайн П. Справочное руководство по конструированию тестов. Киев, 1994.
  30. Классификация и кластер /Под ред. Дж. Вэн Райзина. М. 1980

31. Ковалёва Г.С. О Международной программе PISA-2009 и одном из результатов по критерию: читательская грамотность //Муниципальное образование: инновации и эксперимент, 2010, № 6
32. Коджаспирова Г. М., Коджаспиров А. Ю. Словарь по педагогике. М., Ростов н/Д, 2005.
33. Кондаков Н. И. Логический словарь-справочник. М.: Наука, 1975.
34. Корнилова Т.В., Смирнов С.Д. Методологические основы психологии. СПб: Питер, 2006.
35. Коротков Э.М. Исследование систем управления. М.: ООО «ДеКА», 2000.
36. Кузьмина Н. В. Профессионализм личности преподавателя и мастера производственного обучения. М., 1990.
37. Лаак Я. Психодиагностика: Проблемы содержания и методов. Воронеж: НПО «МОДЭК», 1996.
38. Леонова Е.В. Курсовой проект по дисциплине «Психодиагностика». Методические рекомендации (электронная версия) Обнинск: ИАТЭ, 2006.
39. Леонова Е.В. Математические методы в психологии. Часть 1. Методы первичной обработки психодиагностических данных. Обнинск: ИАТЭ, 2004.
40. Леонова Е.В. Математические методы в психологии. Часть 2. Методы статистической обработки психодиагностических данных. Обнинск: ИАТЭ, 2005.
41. Леонова Е. В. Развитие профессионально важных качеств личности молодого специалиста в области атомной энергетики в период адаптации к обучению в вузе. //«Психология образования: подготовка кадров и психологическое просвещение»: Материалы IV Национальной научно-практической конференции. М., 2007. С. 392-394.
42. Леонова Е. В. Формирование общекультурных компетенций у студентов технического вуза. //Высшее образование в России. 2010. №2, С. 124-131

43. Логинова Н. А., Семенов В. Е. Биографический метод и контент-анализ //Методологические и методические проблемы контент-анализа. – М., Л., 1973. С. 104-106.
44. Майоров А.Н. Тесты школьных достижений: конструирование, проведение, использование. СПб., 1997.
45. Маклаков А.Г. Личностный адаптационный потенциал: его мобилизация и прогнозирование в экстремальных условиях //Психологический журнал, 2001. Т. 22. № 1, С. 16-24.
46. Маклаков А.Г. Общая психология: Учебник для вузов. СПб.: Питер, 2004.
47. Маркова А. К. Психология профессионализма. М.: Знание,1996.
48. Марцинковская Т.Д. Диагностика психического развития детей. М., 1997.
49. Медведев В. Е., Татур Ю. Г. Подготовка преподавателя высшей школы: компетентностный подход. //Высшее образование в России. №11, 2007. С.46-56.
50. Мерлин В. С. Очерк интегрального исследования индивидуальности. М.: Педагогика, 1986.
51. Методические рекомендации по курсовому проектированию для студентов специальности «Психология» /Сост. В.И.Седин. Обнинск: ИАТЭ, 2003.
52. Милерян Е. А. Психология формирования общетрудовых политехнических умений. М., 1973.
53. Министерство образования и науки Российской Федерации. ФГОС высшего профессионального образования. URL: [mon.gov.ru/pro/fgos/vpo/](http://mon.gov.ru/pro/fgos/vpo/) (дата обращения: 15.07.2010)
54. Министерство образования Российской Федерации. Инструктивное письмо № 14-55-996ин/15 «Об активизации самостоятельной работы студентов высших учебных заведений» от 27.11.2002.
55. Наследов А.Д. Математические методы психологического исследования. Анализ и интерпретация данных. Учебное пособие. СПб., 2008.

56. Никандров В.В. Экспериментальная психология. Учебное пособие. СПб.: Издательство «Речь», 2003.
57. Новиков А. М. Методология образования. Издание второе. М.: «Эгвес», 2006.
58. Носс И.Н. Психодиагностика (конспект материалов к практическим занятиям по психодиагностике для экспериментальной психологии). М., 1999.
59. Обозов Н. Н. Межличностные отношения. Л., 1979.
60. Образцов П. И. Методы и методология психологического исследования / П.И. Образцов. Питер, 2004.
61. Общая психодиагностика. //под ред. Бодалёва А.А., Столина В.В. СПб., 2000.
62. Ожегов С. И. Словарь русского языка. М., 1987.
63. Петров В. Ю. Предпосылки развития технического вуза как национального исследовательского университета //Высшее образование в России. 2010. №5.
64. Платонов К. К. Проблемы способностей. М., 1972.
65. Поддьяков А.Н. Конкуренция средств диагностики способностей //Психология способностей: современное состояние и перспективы исследований. М., ИП РАН, 2005. С. 131-136.
66. Поддьяков А.Н. Тестирование интеллекта, конкуренция и рефлексия //Рефлексивные процессы и управление. 2007. № 2, т. 7, С. 46-56
67. Поппер К. Логика социальных наук. //Вопросы философии 1992 № 10, С.65-75.
68. Поппер К. Открытое общество и его враги. т.2. М.: Культурная инициатива, 1992.
69. Практикум по общей, прикладной и экспериментальной психологии //Под ред. А.А.Крылова, С.А.Маничева. С.-Пб., 2000.
70. Практикум по общей, экспериментальной и прикладной психологии: Учебное пособие //Под ред. А.А.Крылова, С.А.Маничева. СПб., 2000.
71. Проективная психология / Пер. с англ. М., 2000.
72. Прыгин Г.С. Введение в психодиагностику. М., 2000.

73. Психологическая диагностика детей и подростков (под ред. К.М.Гуревича, Е.М.Борисовой). М., 1995.
74. Психологическая диагностика детей и подростков //Под ред. К.М.Гуревича, Е.М.Борисовой. М., 1995.
75. Психологическая диагностика: Учеб. для вузов /Под ред. М.К.Акимовой, К.М.Гуревича. СПб.: Питер, 2003.
76. Психологический словарь. /Под ред. В. П. Зинченко, Б. Г. Мещерякова, М., 1996.
77. Реан А.А. Практическая психодиагностика личности. СПб., 2001
78. Равен Дж. Педагогическое тестирование: Проблемы, заблуждения, перспективы / Пер. с англ. М., 2001.
79. Равен Дж. Компетентность в современном обществе. Выявление, развитие и реализация. М., 2002.
80. Рогов Е.И. Настольная книга практического психолога: В 2 кн. М., 2000.
81. Романова Е.С. Графические методы в практической психологии. СПб., 2001.
82. Россолимо Г. И. Психологические профили. Методика. М., 1910.
83. Рубцова Н.Е., Леньков С.Л. Статистические методы в психологии: Учебное пособие. М., 2005.
84. Рукавишников А.А., Соколова М.В. Факторный личностный опросник Кеттелла Р. СПб: «Иматон» 2003.
85. Рыбаков Ф. Е. Атлас для экспериментально-психологического исследования личности с подробным описанием и объяснением таблиц. М., 1910.
86. Рыбников Н. А. Автобиографии рабочих и их изучение. М.; Л., 1930.
87. Сидоренко Е.В. Методы математической обработки в психологии. С.-Пб., 2000.
88. Собчик Л.Н. Методы психологической диагностики. Вып.1–3. М., 1990.



89. Собчик Л.Н. Стандартизированный многофакторный метод исследования личности. СПб., 2000.
90. Степанов С.С. Диагностика интеллекта методом рисуночного теста. М., 1996.
91. Суслов В. И., Чумакова Н.П. Психодиагностика - Учебное пособие. СПб., 1992.
92. Суходольский Г.В. Основы математической статистики для психологов. Л., 1972.
93. Теплов Б.М. О культуре научного исследования //Б.М. Теплов. Избранные труды. Т. 2, М., 1985. С. 310-317.
94. Теплов Б. М. Способность и одаренность Ученые записки Гос. научно-исследов. ин-та психологии, т. 2. М., 1947. С. 3-56.
95. Тимаева С. А. Педагогические условия формирования обобщенных инженерных умений у курсантов младших курсов высших военно-инженерных учебных заведений. Челябинск, 2000
96. Тугаринов В. П. О методах общественных наук. М., 1968.
97. Урбах В.Ю. Статистический анализ в биологических и медицинских исследованиях. М., 1975.
98. Федеральные государственные образовательные стандарты высшего профессионального образования. Сайт Министерства образования и науки РФ. URL: <http://mon.gov.ru/dok/fgos/>
99. Фейерабенд П. Избранные труды по методологии науки. М., 1986.
- 100.Ханин Ю. Л. Стандартный алгоритм адаптации зарубежных опросных методов. – В сб.: Психологические проблемы предсоревновательной подготовки квалифицированных спортсменов. Л., 1977, С. 129-135.
- 101.Холодная М. А. Предисловие к книге Дж. Равена Педагогическое тестирование. М., 2005., С.5-10.
- 102.Чалдини Р. Психология влияния. СПб., 1999.
- 103.Чучалин А. Формирование компетенций выпускников основных образовательных программ //Высшее образование в России, 2008, №12. С.10-18.

- 104.Шадриков В. Д. Проблемы системогенеза профессиональной деятельности. М.: Наука, 1982.
- 105.Шадриков В. Д. Психология деятельности и способности человека. М., 1996.
- 106.Шмелёв А.Г. Основы психодиагностики. М., 1997.
- 107.Шмелев А.Г. Тест как оружие //Психология. Журнал Высшей школы экономики. 2004. № 2. С. 40-53
- 108.Юревич А.В. Психология и методология. М.: ИП РАН, 2005.
- 109.Ядов В. А. Социологическое исследование: Методология. Программа. Методы. М., 1987.
- 110.Allport G. The use of the personal documents in psychological science/ /Soc, Sci. Res Council Bull. 1942. Vol. 49.
- 111.Baldwin A. Personality structure analysis: a statistical method for investigating the single personality //J. abnormal and soc. psychol. 1942 Vol.37.
- 112.Dollard J., Mowrer O. A method of measuring tension in written documents //J. abnormal and soc. psychol. 1947. Vol. 42.
- 113.Lewin K. Hitler youth and Boy Scouts of America: A comparison of aims//Human relations. 1947. Vol. 1.
- 114.Lowenthal L. Biographies in popular magazines //Reader in public opinion and communication / Berelson B., Yanowitz M. Glencoe (Ill.), 1950.
- 115.Osgood Ch., Walker E. Motivation and language behaviour: A content analysis of suicide notes //J. abnormal and soc. psychol. 1959. Vol. 52.
- 116.Sargent S. Emotional stereotypes in the Chicago Tribune //Sociometry. 1939. Vol. 11.
- 117.Seibald H. Studying national character through comparative content analysis//Social forces. 1962. Vol. 40.
- 118.Semenov V. City way of life and problems of intercourse //Psychological problems of the way of life and the life course of personality: (Materials for the Soviet-Finnish symposium). Helsinki, 1984.

119. Stone K, Dunphy D., Smith M., Ogivie D. The general inquirer: A computer approach to content analysis in the behavioural sciences. Cambridge, 1966.
120. Weimer W.B. Psychology and the conceptual foudation of science. Hillsdale, 1976.

## Приложения

## Приложения

### Приложение 1



### ЭТИЧЕСКИЙ КОДЕКС ПСИХОЛОГА

#### *Преамбула*

1. Этический кодекс психолога Российского психологического общества составлен в соответствии с Конституцией Российской Федерации, Федеральным законом Российской Федерации № 152-ФЗ от 27 июля 2006 года «О персональных данных», Уставом Российского психологического общества, Всеобщей декларацией прав человека, Хельсинкской декларацией Всемирной медицинской ассоциации «Этические принципы проведения медицинских исследований с участием людей в качестве субъектов исследования», международной Универсальной декларацией этических принципов для психологов, Этическим метакодексом Европейской федерации психологических ассоциаций.
2. Консультативным и регулирующим органом Российского психологического общества по вопросам профессиональной этики психолога является Этический комитет Российского психологического общества.
3. В настоящем Этическом кодексе термин «Психолог» относится к лицу, имеющему высшее психологическое образование.

4. В настоящем Этическом кодексе термин «Клиент» относится к лицу, группе лиц или организации, которые согласились быть объектом психологических исследований в личных, научных, производственных или социальных интересах или лично обратились к Психологу за психологической помощью.
5. Действие данного Этического кодекса распространяется на все виды деятельности психологов, определенные настоящим Этическим кодексом. Действие данного Этического кодекса распространяется на все формы работы Психолога, в том числе осуществляемые дистанционно или посредством сети Интернет.
6. Профессиональная деятельность психолога характеризуется его особой ответственностью перед клиентами, обществом и психологической наукой, и основана на доверии общества, которое может быть достигнуто только при соблюдении этических принципов профессиональной деятельности и поведения, содержащихся в настоящем Этическом кодексе.
7. Этический кодекс психологов служит: для внутренней регуляции деятельности сообщества психологов; для регуляции отношений психологов с обществом; основой применения санкций при нарушении этических принципов профессиональной деятельности.

## ***I. Этические принципы психолога***

Этика работы психолога основывается на общечеловеческих моральных и нравственных ценностях. Идеалы свободного и всестороннего развития личности и ее уважения, сближения людей, создания справедливого, гуманного, процветающего общества являются определяющими для деятельности психолога. Этические принципы и правила работы психолога формулируют условия, при которых сохраняются и упрочиваются его профессионализм, гуманность его действий, уважение людей, с которыми он работает, и при которых усилия психолога приносят реальную пользу.



## **1. Принцип уважения**

Психолог исходит из уважения личного достоинства, прав и свобод человека, провозглашенных и гарантированных Конституцией Российской Федерации и международными документами о правах человека.

### **Принцип уважения включает:**

#### *1.1. Уважение достоинства, прав и свобод личности*

- (i). Психолог с равным уважением относится к людям вне зависимости от их возраста, пола, сексуальной ориентации, национальности, принадлежности к определенной культуре, этносу и расе, вероисповедания, языка, социально-экономического статуса, физических возможностей и других оснований.
- (ii). Беспристрастность Психолога не допускает предвзятого отношения к Клиенту. Все действия Психолога относительно Клиента должны основываться на данных, полученных научными методами. Субъективное впечатление, которое возникает у Психолога при общении с Клиентом, а также социальное положение Клиента не должны оказывать никакого влияния на выводы и действия Психолога.
- (iii). Психолог избегает деятельности, которая может привести к дискриминации Клиента по любым основаниям.
- (iv). Психологу следует так организовать свою работу, чтобы ни ее процесс, ни ее результаты не наносили вреда здоровью и социальному положению Клиента и связанных с ним лиц.

#### *1.2. Конфиденциальность*

- (i). Информация, полученная Психологом в процессе работы с Клиентом на основе доверительных отношений, не подлежит намеренному или случайному разглашению вне согласованных условий.
- (ii). Результаты исследования должны быть представлены таким образом, чтобы они не могли скомпрометировать Клиента, Психолога или психологическую науку.
- (iii). Психодиагностические данные студентов, полученные при их обучении, должны рассматриваться конфиден-

циально. Сведения о Клиентах также должны рассматриваться конфиденциально.

- (iv). Демонстрируя конкретные случаи своей работы, Психолог должен обеспечить защиту достоинства и благополучия Клиента.
- (v). Психолог не должен отыскивать о Клиенте информацию, которая выходит за рамки профессиональных задач Психолога.
- (vi). Клиент имеет право на консультацию Психолога или работу с ним без присутствия третьих лиц.
- (vii). Неконтролируемое хранение данных, полученных при исследованиях, может нанести вред Клиенту, Психологу и обществу в целом. Порядок обращения с полученными в исследованиях данными и порядок их хранения должны быть жестко регламентированы.

### *1.3. Осведомленность и добровольное согласие Клиента*

- (i). Клиент должен быть извещен о цели работы, о применяемых методах и способах использования полученной информации. Работа с Клиентом допускается только после того, как Клиент дал информированное согласие в ней участвовать. В случае, если Клиент не в состоянии сам принимать решение о своем участии в работе, такое решение должно быть принято его законными представителями.
- (ii). Психолог должен сообщать Клиенту обо всех основных шагах или лечебных действиях. В случае стационарного лечения Психолог должен информировать Клиента о возможных рисках и об альтернативных методах лечения, включая непсихологические.
- (iii). Видео- или аудиозаписи консультации или лечения Психолог может делать только после того, как получит согласие на это со стороны Клиента. Это положение распространяется и на телефонные переговоры. Ознакомление третьих лиц с видео-, аудиозаписями консультации и телефонными переговорами Психолог может разрешить только после получения согласия на это со стороны Клиента.

- (iv). Участие в психологических экспериментах и исследованиях должно быть добровольным. Клиент должен быть проинформирован в понятной для него форме о целях, особенностях исследования и возможном риске, дискомфорте или нежелательных последствиях, чтобы он мог самостоятельно принять решение о сотрудничестве с Психологом. Психолог обязан предварительно удостовериться в том, что достоинство и личность Клиента не пострадают. Психолог должен принять все необходимые предосторожности для обеспечения безопасности и благополучия Клиента и сведения к минимуму возможности непредвиденного риска.
- (v). В тех случаях, когда предварительное исчерпывающее раскрытие информации противоречит задачам проводимого исследования, Психолог должен принять специальные меры предосторожности для обеспечения благополучия испытуемых. В тех случаях, когда это возможно, и при условии, что сообщаемая информация не нанесет вреда Клиенту, все разъяснения должны быть сделаны после окончания эксперимента.

#### *1.4. Самоопределение Клиента*

- (i). Психолог признает право Клиента на сохранение максимальной автономии и самоопределения, включая общее право вступать в профессиональные отношения с психологом и прекращать их.
- (ii). Клиентом может быть любой человек в случае своей несомненной дееспособности по возрасту, состоянию здоровья, умственному развитию, физической независимости. В случае недостаточной дееспособности человека решение о его сотрудничестве с Психологом принимает лицо, представляющее интересы этого человека по закону.
- (iii). Психолог не должен препятствовать желанию Клиента привлечь для консультации другого психолога (в тех случаях, когда к этому нет юридических противопоказаний).

## **2. Принцип компетентности**

Психолог должен стремиться обеспечивать и поддерживать высокий уровень компетентности в своей работе, а также признавать границы своей компетентности и своего опыта. Психолог должен предоставлять только те услуги и использовать только те методы, которым обучался и в которых имеет опыт.

### **Принцип компетентности включает:**

#### *2.1. Знание профессиональной этики*

- (i). Психолог должен обладать исчерпывающими знаниями в области профессиональной этики и обязан знать положения настоящего Этического кодекса. В своей работе Психолог должен руководствоваться этическими принципами.
- (ii). Если персонал или студенты выступают в качестве экспериментаторов в проведении психодиагностических процедур, Психолог должен обеспечить, независимо от их собственной ответственности, соответствие совершаемых ими действий профессиональным требованиям.
- (iii). Психолог несет ответственность за соответствие профессионального уровня персонала, которым он руководит, требованиям выполняемой работы и настоящего Этического кодекса.
- (iv). В своих рабочих контактах с представителями других профессий Психолог должен проявлять лояльность, терпимость и готовность помочь.

#### *2.2. Ограничения профессиональной компетентности*

- (i). Психолог обязан осуществлять практическую деятельность в рамках собственной компетентности, основанной на полученном образовании и опыте.
- (ii). Только Психолог осуществляет непосредственную (анкетирование, интервьюирование, тестирование, электрофизиологическое исследование, психотерапия, тренинг и др.) или опосредованную (биографический метод, метод наблюдения, изучение продуктов деятельности Клиента и др.) работу с Клиентом.
- (iii). Психолог должен владеть методами психодиагностической беседы, наблюдения, психолого-педагогического

воздействия на уровне, достаточном, чтобы поддерживать у Клиента чувство симпатии, доверия и удовлетворения от общения с Психологом.

- (iv). Если Клиент болен, то работа с ним допустима только с разрешения врача или согласия других лиц, представляющих интересы Клиента.

### *2.3. Ограничения применяемых средств*

- (i). Психолог может применять методики, которые адекватны целям проводимого исследования, возрасту, полу, образованию, состоянию Клиента, условиям эксперимента. Психодиагностические методики, кроме этого, обязательно должны быть стандартизованными, нормализованными, надежными, валидными и адаптированными к контингенту испытуемых.
- (ii). Психолог должен применять методы обработки и интерпретации данных, получившие научное признание. Выбор методов не должен определяться научными пристрастиями Психолога, его общественными увлечениями, личными симпатиями к Клиентам определенного типа, социального положения или профессиональной деятельности.
- (iii). Психологу запрещается представлять в результатах исследования намеренно искаженные первичные данные, заведомо ложную и некорректную информацию. В случае обнаружения Психологом существенной ошибки в своем исследовании после того, как исследование было опубликовано, он должен предпринять все возможные действия по исправлению ошибки и дальнейшему опубликованию исправлений.

### *2.4. Профессиональное развитие*

- (i). Психолог должен постоянно повышать уровень своей профессиональной компетентности и свою осведомленность в области этики психологической работы (исследования).

### *2.5. Невозможность профессиональной деятельности в определенных условиях*

- (i). Если какие-либо обстоятельства вынуждают Психолога преждевременно прекратить работу с Клиентом и это

может отрицательно сказаться на состоянии Клиента, Психолог должен обеспечить продолжение работы с Клиентом.

- (ii). Психолог не должен выполнять свою профессиональную деятельность в случае, когда его способности или суждения находятся под неблагоприятным воздействием.

### **3. Принцип ответственности**

Психолог должен помнить о своих профессиональных и научных обязательствах перед своими клиентами, перед профессиональным сообществом и обществом в целом. Психолог должен стремиться избегать причинения вреда, должен нести ответственность за свои действия, а также гарантировать, насколько это возможно, что его услуги не являются злоупотреблением.

#### **Принцип ответственности включает:**

##### *3.1. Основная ответственность*

- (i). Решение Психолога осуществить исследовательский проект или вмешательство предполагает его ответственность за возможные научные и социальные последствия, включая воздействие на лиц, группы и организации, участвующие в исследовании или вмешательстве, а также не прямой эффект, как, например, влияние научной психологии на общественное мнение и на развитие представлений о социальных ценностях.
- (ii). Психолог должен осознавать специфику взаимодействия с Клиентом и вытекающую из этого ответственность. Ответственность особенно велика в случае, если в качестве испытуемых или клиентов выступают лица, страдающие от медикаментозной зависимости, или лица, ограниченные в своих действиях, а также, если программа исследования или вмешательства целенаправленно ограничивает дееспособность Клиента.
- (iii). Если Психолог приходит к заключению, что его действия не приведут к улучшению состояния Клиента или представляют риск для Клиента, он должен прекратить вмешательство.

##### *3.2. Ненанесение вреда*

- (i). Психолог применяет только такие методики исследования или вмешательства, которые не являются опасными



для здоровья, состояния Клиента, не представляют Клиента в результатах исследования в ложном, искаженном свете, и не дают сведений о тех психологических свойствах и особенностях Клиента, которые не имеют отношения к конкретным и согласованным задачам психологического исследования.

### *3.3. Решение этических дилемм*

- (i). Психолог должен осознавать возможность возникновения этических дилемм и нести свою персональную ответственность за их решение. Психологи консультируются по этим вопросам со своими коллегами и другими значимыми лицами, а также информируют их о принципах, отраженных в Этическом кодексе.
- (ii). В случае, если у Психолога в связи с его работой возникли вопросы этического характера, он должен обратиться в Этический комитет Российского психологического общества за консультацией.

## **4. Принцип честности**

Психолог должен стремиться содействовать открытости науки, обучения и практики в психологии. В этой деятельности психолог должен быть честным, справедливым и уважающим своих коллег. Психологу надлежит четко представлять свои профессиональные задачи и соответствующие этим задачам функции.

### **Принцип честности включает:**

#### *4.1. Осознание границ личных и профессиональных возможностей*

- (i). Психолог должен осознавать ограниченность как своих возможностей, так и возможностей своей профессии. Это условие установления диалога между профессионалами различных специальностей.

#### *4.2. Честность*

- (i). Психолог и Клиент (или сторона, инициирующая и оплачивающая психологические услуги для Клиента) до заключения соглашения оговаривают вопросы вознаграждения и иные существенные условия работы, такие как распределение прав и обязанностей между Психологом и Клиентом (или стороной, оплачивающей психологические услуги) или процедура хранения и при-

менения результатов исследования. Психолог должен известить Клиента или работодателя о том, что его деятельность в первую очередь подчиняется профессиональным, а не коммерческим принципам.

При приеме на работу Психолог должен поставить своего работодателя в известность о том, что:

- в пределах своей компетенции он будет действовать независимо;
- он обязан соблюдать принцип конфиденциальности: этого требует закон;
- профессиональное руководство его работой может осуществлять только психолог;
- для него невозможно выполнение непрофессиональных требований или требований, нарушающих данный Этический кодекс.

При приеме Психолога на работу работодатель должен получить текст данного Этического кодекса.

- (ii). Публичное распространение сведений об оказываемых Психологом услугах служит целям принятия потенциальными Клиентами информированного решения о вступлении в профессиональные отношения с Психологом. Подобная реклама приемлема только в том случае, если она не содержит ложных или искаженных сведений, отражает объективную информацию о предоставляемых услугах и отвечает правилам приличия.
- (iii). Психологу запрещается организовывать рекламу себе или какому-либо определенному методу вмешательства или лечения. Реклама в целях конкуренции ни при каких условиях не должна обманывать потенциальных Клиентов. Психолог не должен преувеличивать эффективность своих услуг, делать заявлений о превосходстве своих профессиональных навыков и применяемых методик, а также давать гарантии результативности оказываемых услуг.
- (iv). Психологу не разрешается предлагать скидку или вознаграждение за направление к ним нему Клиентов или заключать соглашения с третьими лицами с этой целью.

#### *4.3. Прямота и открытость*

- (i). Психолог должен нести ответственность за предоставляемую им информацию и избегать ее искажения в исследовательской и практической работе.
- (ii). Психолог формулирует результаты исследования в терминах и понятиях, принятых в психологической науке, подтверждая свои выводы предъявлением первичных материалов исследования, их математико-статистической обработкой и положительным заключением компетентных коллег. При решении любых психологических задач проводится исследование, всегда опирающееся на предварительный анализ литературных данных по поставленному вопросу.
- (iii). В случае возникновения искажения информации психолог должен проинформировать об этом участников взаимодействия и заново установить степень доверия.

#### *4.4. Избегание конфликта интересов*

- (i). Психолог должен осознавать проблемы, которые могут возникнуть в результате двойственных отношений. Психолог должен стараться избегать отношений, которые приводят к конфликтам интересов или эксплуатации отношений с Клиентом в личных интересах.
- (ii). Психолог не должен использовать профессиональные отношения в личных, религиозных, политических или идеологических интересах.
- (iii). Психолог должен осознавать, что конфликт интересов может возникнуть после формального прекращения отношений Психолога с Клиентом. Психолог в этом случае также несет профессиональную ответственность.
- (iv). Психолог не должен вступать в какие бы то ни было личные отношения со своими Клиентами.

#### *4.5. Ответственность и открытость перед профессиональным сообществом*

- (i). Результаты психологических исследований должны быть доступны для научной общественности. Возможность неверной интерпретации должна быть предупреждена корректным, полным и недвусмысленным изложением.

Данные об участниках эксперимента должны быть анонимными. Дискуссии и критика в научных кругах служат развитию науки и им не следует препятствовать.

- (ii). Психолог обязан уважать своих коллег и не должен необъективно критиковать их профессиональные действия.
- (iii). Психолог не должен своими действиями способствовать вытеснению коллеги из его сферы деятельности или лишению его работы.
- (iv). Если Психолог считает, что его коллега действует непрофессионально, он должен указать ему на это конфиденциально.

## ***II. Нарушение Этического кодекса психолога***

1. Нарушение Этического кодекса психолога включает в себя игнорирование изложенных в нем положений, неверное их толкование или намеренное нарушение. Нарушение Этического кодекса может стать предметом жалобы.

2. Жалоба на нарушение Этического кодекса психолога может быть подана в Этический комитет Российского психологического общества в письменном виде любым физическим и юридическим лицом. Рассмотрение жалоб и вынесение решений по ним осуществляется в установленном порядке Этическим комитетом Российского психологического общества.

3. В качестве санкций, применяемых к Психологу, нарушившему Этический кодекс, могут выступать: предупреждение от имени Российского психологического общества (общественное порицание), приостановление членства в Российском психологическом обществе, сопровождающееся широким информированием общественности и потенциальных клиентов об исключении данного специалиста из действующего реестра психологов РПО. Информация о применяемых санкциях является общедоступной и передается в профессиональные психологические ассоциации других стран.

4. В случае серьезных нарушений Этического кодекса Российское психологическое общество может ходатайствовать о привлечении Психолога к суду.

*Настоящий Этический кодекс психолога принят 14 февраля 2012 года V съездом Российского психологического общества.*

### ***Клятва российского психолога***

Получая высокое звание российского психолога и приступая к профессиональной деятельности, я торжественно клянусь:

- честно исполнять свой профессиональный долг, соблюдая этические принципы работы российского психолога;
- посвятить свои знания и умения сохранению и укреплению психологического здоровья и благополучия человека;
- быть всегда готовым оказать психологическую помощь;
- проявлять высочайшее уважение к жизни, правам, благополучию и достоинству человека;
- хранить благодарность и уважение к своим учителям;
- быть требовательным и справедливым к своим ученикам, способствовать их профессиональному росту;
- доброжелательно относиться к коллегам;
- обращаться к коллегам за помощью и советом, если этого требуют интересы людей, обратившихся ко мне за помощью;
- никогда не отказывать коллегам в помощи и совете;
- постоянно совершенствовать свое профессиональное мастерство;
- беречь и развивать славные и благородные традиции российской психологической науки.

Да буду я чист в своих делах и помыслах, да не нарушу эту клятву вовек.

*Клятва торжественно приносится при вручении диплома о высшем психологическом образовании.*

## Приложение 2

### Статистические таблицы

**Таблица 1**

**Критические значения t-критерия Стьюдента**  
**для уровней статистической значимости  $p=0,05$ ,  $p=0,01$  и  $p=0,001$**   
**и числа степеней свободы  $k$**

k k	p			k	p		
	0,05	0,01	0,001		0,05	0,01	0,001
1	12,71	63,66	64,60	18	2,10	2,88	3,92
2	4,30	9,92	31,60	19	2,09	2,86	3,88
3	3,18	5,84	12,92	20	2,09	2,85	3,85
4	2,78	4,60	8,61	21	2,08	2,83	3,82
5	2,57	4,03	6,87	22	2,07	2,82	3,79
6	2,45	3,71	5,96	23	2,07	2,81	3,77
7	2,37	3,50	5,41	24	2,06	2,80	3,75
8	2,31	3,36	5,04	25	2,06	2,79	3,73
9	2,26	3,25	4,78	26	2,06	2,78	3,71
10	2,23	3,17	4,59	27	2,05	2,77	3,69
11	2,20	3,11	4,44	28	2,05	2,76	3,67
12	2,18	3,05	4,32	29	2,05	2,76	3,66
13	2,16	3,01	4,22	30	2,04	2,75	3,65
14	2,14	2,98	4,14	40	2,02	2,70	3,55
15	2,13	2,95	4,07	60	2,00	2,66	3,46
16	2,12	2,92	4,02	120	1,98	2,62	3,37
17	2,11	2,90	3,97	$\infty$	1,96	2,58	3,29



**Таблица 2**

**Критические значения F-критерия Фишера**

**для уровней статистической значимости  $p=0,05$  и  $p=0,01$**

**и числа степеней свободы  $df_1$  – в числителе,  $df_2$  – в знаменателе**

$df_1$	1	2	3	4	5	6	7	8	9	10	11	12
$df_2$	<b><math>p=0,05</math></b>											
1	161	200	216	225	230	234	237	239	241	242	243	244
2	18,51	19,00	19,16	19,25	19,3	19,33	19,36	19,37	19,38	19,39	19,40	19,41
3	10,13	9,55	9,28	9,12	9,01	8,94	8,88	8,84	8,81	8,78	8,76	8,74
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,93	5,91
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,78	4,74	4,70	4,68
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,03	4,00
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,63	3,60	3,57
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,34	3,31	3,28
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,13	3,10	3,07
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,97	2,94	2,91
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,86	2,82	2,79
12	4,75	3,88	3,49	3,26	3,11	3,00	2,92	2,85	2,80	2,76	2,72	2,69
13	4,67	3,80	3,41	3,18	3,02	2,92	2,84	2,77	2,72	2,67	2,63	2,60
14	4,60	3,74	3,34	3,11	2,96	2,85	2,77	2,70	2,65	2,60	2,56	2,53
15	4,54	3,68	3,29	3,06	2,90	2,79	2,70	2,64	2,59	2,55	2,51	2,48
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,45	2,42
$df_2$	<b><math>p=0,01</math></b>											
1	4052	4999	5403	5625	5764	5859	5928	5981	6022	6056	6082	6106
2	98,49	99,00	99,17	99,25	99,30	99,33	99,36	99,37	99,39	99,40	99,41	99,42
3	34,12	30,82	29,46	28,71	28,24	27,91	27,67	27,49	27,34	27,23	27,13	27,05
4	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80	14,66	14,54	14,45	14,37
5	16,26	13,27	12,06	11,39	10,97	10,67	10,45	10,29	10,15	10,05	9,96	9,89
6	13,74	10,92	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87	7,79	7,72
7	12,25	9,55	8,45	7,85	7,46	7,19	7,00	6,84	6,71	6,62	6,54	6,47
8	11,26	8,65	7,59	7,01	6,63	6,37	6,19	6,03	5,91	5,82	5,74	5,67
9	10,56	8,02	6,99	6,42	6,06	5,80	5,62	5,47	5,35	5,26	5,18	5,11
10	10,04	7,56	6,55	5,99	5,64	5,39	5,21	5,06	4,95	4,85	4,78	4,71
11	9,65	7,20	6,22	5,67	5,32	5,07	4,88	4,74	4,63	4,54	4,46	4,40
12	9,33	6,93	5,95	5,41	5,06	4,82	4,65	4,50	4,39	4,30	4,22	4,16
13	9,07	6,70	5,74	5,20	4,86	4,62	4,44	4,30	4,19	4,10	4,02	3,96
14	8,86	6,51	5,56	5,03	4,69	4,46	4,28	4,14	4,03	3,94	3,86	3,80
15	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89	3,80	3,73	3,67
16	8,53	6,23	5,29	4,77	4,44	4,20	4,03	3,89	3,78	3,69	3,61	3,55

df <sub>1</sub>	1	2	3	4	5	6	7	8	9	10	11	12
df <sub>2</sub>	<b>p= 0,05</b>											
17	4,45	3,59	3,20	2,96	2,81	2,70	2,62	2,55	2,50	2,45	2,41	2,38
18	4,41	3,55	3,16	2,93	2,77	3,66	2,58	2,51	2,46	2,41	2,37	2,34
19	4,38	3,52	3,13	2,90	2,74	2,63	2,55	2,48	2,43	2,36	2,34	2,31
20	4,35	3,49	3,10	2,87	2,71	2,60	2,52	2,45	2,40	2,35	2,31	2,28
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,28	2,25
22	4,30	3,44	3,05	2,82	2,66	2,55	2,47	2,40	2,35	2,30	2,26	2,23
23	4,28	3,42	3,03	2,80	2,64	2,53	2,45	2,38	2,32	2,28	2,24	2,20
24	4,26	3,40	3,01	2,78	2,62	2,51	2,43	2,36	2,30	2,26	2,22	2,18
25	4,24	3,38	2,99	2,76	2,60	2,49	2,41	2,34	2,28	2,24	2,20	2,16
26	4,22	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22	2,18	2,15
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,30	2,25	2,20	2,16	2,13
28	4,20	3,34	2,95	2,71	2,56	2,44	2,36	2,29	2,24	2,19	2,15	2,12
29	4,18	3,33	2,93	2,70	2,54	2,43	2,35	2,28	2,22	2,18	2,14	2,10
30	4,17	3,32	2,92	2,69	2,53	2,42	2,34	2,27	2,21	2,16	2,12	2,09
32	4,15	3,30	2,90	2,67	2,51	2,40	2,32	2,25	2,19	2,14	2,10	2,07
34	4,13	3,28	2,88	2,65	2,49	2,38	2,30	2,23	2,17	2,12	2,08	2,05
<b>p= 0,01</b>												
17	8,40	6,11	5,18	4,67	4,34	4,10	3,93	3,79	3,68	3,59	3,52	3,45
18	8,28	6,01	5,09	4,58	4,25	4,01	3,85	3,71	3,60	3,51	3,44	3,37
19	8,18	5,93	5,01	4,50	4,17	3,94	3,77	3,63	3,52	3,43	3,36	3,30
20	8,10	5,85	4,94	4,43	4,10	3,87	3,71	3,56	3,45	3,37	3,30	3,23
21	8,02	5,78	4,87	4,37	4,04	3,81	3,65	3,51	3,40	3,31	3,24	3,17
22	7,94	5,72	4,82	4,31	3,99	3,76	3,59	3,45	3,35	3,26	3,18	3,12
23	7,88	5,66	4,76	4,26	3,94	3,71	3,54	3,41	3,30	3,21	3,14	3,07
24	7,82	5,61	4,72	4,22	3,90	3,67	3,50	3,36	3,25	3,17	3,09	3,03
25	7,77	5,57	4,68	4,18	3,86	3,63	3,46	3,32	3,21	3,13	3,05	2,99
26	7,72	5,53	4,64	4,14	3,82	3,59	3,42	3,29	3,17	3,09	3,02	2,96
27	7,68	5,49	4,60	4,11	3,79	3,56	3,39	3,26	3,14	3,06	2,98	2,93
28	7,64	5,45	4,57	4,07	3,76	3,53	3,36	3,23	3,11	3,03	2,95	2,90
29	7,60	5,42	4,54	4,04	3,73	3,50	3,33	3,20	3,08	3,00	2,92	2,87
30	7,56	5,39	4,51	4,02	3,70	3,47	3,30	3,17	3,06	2,98	2,90	2,84
32	7,50	5,34	4,46	3,97	3,66	3,42	3,25	3,12	3,01	2,94	2,86	2,80
34	7,44	5,29	4,42	3,93	3,61	3,38	3,21	3,08	2,97	2,89	2,82	2,76

df <sub>1</sub>	14	16	20	24	30	40	50	75	100	200	500	∞
df <sub>2</sub>	<b>p= 0,05</b>											
1	245	246	248	249	250	251	252	253	253	254	254	254
2	19,42	19,43	19,44	19,45	19,46	19,47	19,47	19,48	19,49	19,49	19,50	19,50
3	8,71	8,69	8,66	8,64	8,62	8,60	8,58	8,57	8,56	8,54	8,54	8,53
4	5,87	5,84	5,80	5,77	5,74	5,71	5,70	5,68	5,66	5,65	5,64	5,63
5	4,64	4,60	4,56	4,53	4,50	4,46	4,44	4,42	4,40	4,38	4,37	4,36
6	3,96	3,92	3,87	3,84	3,81	3,77	3,75	3,72	3,71	3,69	3,68	3,67
7	3,52	3,49	3,44	3,41	3,38	3,34	3,32	3,29	3,28	3,25	3,24	3,23
8	3,23	3,20	3,15	3,12	3,08	3,05	3,03	3,00	2,98	2,96	2,94	2,93
9	3,02	2,98	2,93	2,90	2,86	2,82	2,89	2,77	2,76	2,73	2,72	2,71
10	2,86	2,82	2,77	2,74	2,70	2,67	2,64	2,61	2,59	2,56	2,55	2,54
11	2,74	2,70	2,65	2,61	2,57	2,53	2,50	2,47	2,45	2,42	2,41	2,40
12	2,64	2,60	2,54	2,50	2,46	2,42	2,40	2,36	2,35	2,32	2,31	2,30
13	2,55	2,51	2,46	2,42	2,38	2,34	2,32	2,28	2,26	2,24	2,22	2,21
14	2,48	2,44	2,39	2,35	2,31	2,27	2,24	2,21	2,19	2,16	2,14	2,13
15	2,43	2,39	2,33	2,29	2,25	2,21	2,18	2,15	2,12	2,10	2,08	2,07
16	2,37	2,33	2,28	2,24	2,20	2,16	2,13	2,09	2,07	2,04	2,02	2,01
	<b>p= 0,01</b>											
1	6142	6169	6208	6234	6261	6286	6302	6323	6334	6352	6361	6366
2	99,43	99,44	99,45	99,46	99,47	99,48	99,48	99,4	99,49	99,49	99,50	99,50
3	26,92	26,83	26,69	26,60	26,50	26,41	26,35	26,27	26,23	26,18	26,14	26,12
4	14,24	14,15	14,02	13,93	13,83	13,74	13,69	13,61	13,57	13,52	13,48	13,46
5	9,77	9,68	9,55	9,47	9,38	9,29	9,24	9,17	9,13	9,07	9,04	9,02
6	7,60	7,52	7,39	7,31	7,23	7,14	7,09	7,02	6,99	6,94	6,90	6,88
7	6,35	6,27	6,15	6,07	5,98	5,90	5,85	5,78	5,75	5,70	5,67	5,65
8	5,56	5,48	5,36	5,28	5,20	5,11	5,06	5,00	4,96	4,91	4,88	4,86
9	5,00	4,92	4,80	4,73	4,64	4,56	4,51	4,45	4,41	4,36	4,33	4,31
10	4,60	4,52	4,41	4,33	4,25	4,17	4,12	4,05	4,01	3,96	3,93	3,91
11	4,29	4,21	4,10	4,02	3,94	3,86	3,80	3,74	3,70	3,66	3,62	3,60
12	4,05	3,98	3,86	3,78	3,70	3,61	3,56	3,49	3,46	3,41	3,38	3,36
13	3,85	3,78	3,67	3,59	3,51	3,42	3,37	3,30	3,27	3,21	3,18	3,16
14	3,70	3,62	3,51	3,43	3,34	3,26	3,21	3,14	3,11	3,06	3,02	3,00
15	3,56	3,48	3,36	3,29	3,20	3,12	3,07	3,00	2,97	2,92	2,89	2,87
16	3,45	3,37	3,25	3,18	3,10	3,01	2,96	2,98	2,86	2,80	2,77	2,75

df <sub>1</sub>	14	16	20	24	30	40	50	75	100	200	500	∞
df <sub>2</sub>	<b>p= 0,05</b>											
17	2,33	2,29	2,23	2,19	2,15	2,11	2,08	2,04	2,02	1,99	1,97	1,96
18	2,29	2,25	2,19	2,15	2,11	2,07	2,04	2,00	1,98	1,95	1,93	1,92
19	2,26	2,21	2,15	2,11	2,07	2,02	2,00	1,96	1,94	1,91	1,90	1,88
20	2,23	2,18	2,12	2,08	2,04	1,99	1,96	1,92	1,90	1,87	1,85	1,84
21	2,20	2,15	2,09	2,05	2,00	1,96	1,93	1,89	1,87	1,84	1,82	1,81
22	2,18	2,13	2,07	2,03	1,98	1,93	1,91	1,87	1,84	1,81	1,80	1,78
23	2,14	2,10	2,04	2,00	1,96	1,91	1,88	1,84	1,82	1,79	1,77	1,76
24	2,13	2,09	2,02	1,98	1,94	1,89	1,86	1,82	1,80	1,76	1,74	1,73
25	2,11	2,06	2,00	1,96	1,92	1,87	1,84	1,80	1,77	1,74	1,72	1,71
26	2,10	2,05	1,99	1,95	1,90	1,85	1,82	1,78	1,76	1,72	1,70	1,69
27	2,08	2,03	1,97	1,93	1,88	1,84	1,80	1,76	1,74	1,71	1,68	1,67
28	2,06	2,02	1,96	1,91	1,87	1,81	1,78	1,75	1,72	1,69	1,67	1,65
29	2,05	2,00	1,94	1,90	1,85	1,80	1,77	1,73	1,71	1,68	1,65	1,64
30	2,04	1,99	1,93	1,89	1,84	1,79	1,76	1,72	1,69	1,66	1,64	1,62
32	2,02	1,97	1,91	1,86	1,82	1,76	1,74	1,69	1,67	1,64	1,61	1,59
34	2,00	1,95	1,89	1,84	1,80	1,74	1,71	1,67	1,64	1,61	1,59	1,57
	<b>p= 0,01</b>											
17	3,35	3,27	3,16	3,08	3,00	2,92	2,86	2,79	2,76	2,70	2,67	2,65
18	3,27	3,19	3,07	3,00	2,91	2,83	2,78	2,71	2,68	2,62	2,59	2,57
19	3,19	3,12	3,00	2,92	2,84	2,76	2,70	2,63	2,60	2,54	2,51	2,49
20	3,13	3,05	2,94	2,86	2,77	2,69	2,63	2,56	2,53	2,47	2,44	2,42
21	3,07	2,99	2,88	2,80	2,72	2,63	2,58	2,51	2,47	2,42	2,38	2,36
22	3,02	2,94	2,83	2,75	2,67	2,58	2,53	2,46	2,42	2,37	2,33	2,31
23	2,97	2,89	2,78	2,70	2,62	2,53	2,48	2,41	2,37	2,32	2,28	2,26
24	2,93	2,85	2,74	2,66	2,58	2,49	2,44	2,36	2,33	2,27	2,23	2,21
25	2,89	2,81	2,70	2,62	2,54	2,45	2,40	2,32	2,29	2,23	2,19	2,17
26	2,86	2,77	2,66	2,58	2,50	2,41	2,36	2,28	2,25	2,19	2,15	2,13
27	2,83	2,74	2,63	2,55	2,47	2,38	2,33	2,25	2,21	2,16	2,12	2,10
28	2,80	2,71	2,60	2,52	2,44	2,35	2,30	2,22	2,18	2,13	2,09	2,06
29	2,77	2,68	2,57	2,49	2,41	2,32	2,27	2,19	2,15	2,10	2,06	2,03
30	2,74	2,66	2,55	2,47	2,38	2,29	2,24	2,16	2,13	2,07	2,03	2,01
32	2,70	2,62	2,51	2,42	2,34	2,25	2,20	2,12	2,08	2,02	1,98	1,96
34	2,66	2,58	2,47	2,38	2,30	2,21	2,15	2,08	2,04	1,98	1,94	1,91

df <sub>1</sub>	14	16	20	24	30	40	50	75	100	200	500	∞
df <sub>2</sub>	<b>p= 0,05</b>											
36	1,98	1,93	1,87	1,82	1,78	1,72	1,69	1,65	1,62	1,59	1,56	1,55
38	1,96	1,92	1,85	1,80	1,76	1,71	1,67	1,63	1,60	1,57	1,54	1,53
40	1,95	1,90	1,84	1,79	1,74	1,69	1,66	1,61	1,59	1,55	1,53	1,51
42	1,94	1,89	1,82	1,78	1,73	1,68	1,64	1,60	1,57	1,54	1,51	1,49
44	1,92	1,88	1,81	1,76	1,72	1,66	1,63	1,58	1,56	1,52	1,50	1,48
46	1,91	1,87	1,80	1,75	1,71	1,65	1,62	1,57	1,54	1,51	1,48	1,46
48	1,90	1,86	1,79	1,74	1,70	1,64	1,61	1,56	1,53	1,50	1,47	1,45
50	1,90	1,85	1,78	1,74	1,69	1,63	1,60	1,55	1,52	1,48	1,46	1,44
55	1,88	1,83	1,76	1,72	1,67	1,61	1,58	1,52	1,50	1,46	1,43	1,41
60	1,86	1,81	1,75	1,70	1,65	1,59	1,56	1,50	1,48	1,44	1,41	1,39
65	1,85	1,80	1,73	1,68	1,63	1,57	1,54	1,49	1,46	1,42	1,39	1,37
70	1,84	1,79	1,72	1,67	1,62	1,56	1,53	1,47	1,45	1,40	1,37	1,35
80	1,82	1,77	1,70	1,65	1,60	1,54	1,51	1,45	1,42	1,38	1,35	1,32
100	1,79	1,75	1,68	1,63	1,57	1,51	1,48	1,42	1,39	1,34	1,30	1,28
125	1,77	1,72	1,65	1,60	1,55	1,49	1,45	1,39	1,36	1,31	1,27	1,25
150	1,76	1,71	1,64	1,59	1,54	1,47	1,44	1,37	1,34	1,29	1,25	1,22
200	1,74	1,69	1,62	1,57	1,52	1,45	1,42	1,35	1,32	1,26	1,22	1,19
400	1,72	1,67	1,60	1,54	1,49	1,42	1,38	1,32	1,28	1,22	1,16	1,13
1000	1,70	1,65	1,58	1,53	1,47	1,41	1,36	1,30	1,26	1,19	1,13	1,08
∞	1,69	1,64	1,57	1,52	1,46	1,40	1,35	1,28	1,24	1,17	1,11	1,00
	<b>p= 0,01</b>											
36	2,62	2,54	2,43	2,35	2,26	2,17	2,12	2,04	2,00	1,94	1,90	1,87
38	2,59	2,51	2,40	2,32	2,22	2,14	2,08	2,00	1,97	1,90	1,86	1,84
40	2,56	2,49	2,37	2,29	2,20	2,11	2,05	1,97	1,94	1,88	1,84	1,81
42	2,54	2,46	2,35	2,26	2,17	2,08	2,02	1,94	1,91	1,85	1,80	1,78
44	2,52	2,44	2,32	2,24	2,15	2,06	2,00	1,92	1,88	1,82	1,78	1,75
46	2,50	2,42	2,30	2,22	2,13	2,04	1,98	1,90	1,86	1,80	1,76	1,72
48	2,48	2,40	2,28	2,20	2,11	2,02	1,96	1,88	1,84	1,78	1,73	1,70
50	2,46	2,39	2,26	2,18	2,10	2,00	1,94	1,86	1,82	1,76	1,71	1,68
55	2,43	2,35	2,23	2,15	2,06	1,96	1,90	1,82	1,78	1,71	1,66	1,64
60	2,40	2,32	2,20	2,12	2,03	1,93	1,87	1,79	1,74	1,68	1,63	1,60
65	2,37	2,30	2,18	2,09	2,00	1,90	1,84	1,76	1,71	1,64	1,60	1,56
70	2,35	2,28	2,15	2,07	1,98	1,88	1,82	1,74	1,69	1,62	1,56	1,53
80	2,32	2,24	2,11	2,03	1,94	1,84	1,78	1,70	1,65	1,57	1,52	1,49
100	2,26	2,19	2,06	1,98	1,89	1,79	1,73	1,64	1,59	1,51	1,46	1,43
125	2,23	2,15	2,03	1,94	1,85	1,75	1,68	1,59	1,54	1,46	1,40	1,37
150	2,20	2,12	2,00	1,91	1,83	1,72	1,66	1,56	1,51	1,43	1,37	1,33
200	2,17	2,09	1,97	1,88	1,79	1,69	1,62	1,53	1,48	1,39	1,33	1,28
400	2,12	2,04	1,92	1,84	1,74	1,64	1,57	1,47	1,42	1,32	1,24	1,19
1000	2,09	2,01	1,89	1,81	1,71	1,61	1,54	1,44	1,38	1,28	1,19	1,11
∞	2,07	1,99	1,87	1,79	1,69	1,59	1,52	1,41	1,36	1,25	1,15	1,00

**Таблица 3**  
**Критические значения  $U$ -критерия Манна-Уитни для уровней**  
**статистической значимости  $p=0,05$  и  $p=0,01$  и объёмов**  
**сопоставляемых выборок  $2 \leq n_1 \leq 20, 3 \leq n_2 \leq 20$**

$p=0.05$

$n_1 \backslash n_2$	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
3	-	0																	
4	-	0	1																
5	0	1	2	4															
6	0	2	3	5	7														
7	0	2	4	6	8	11													
8	1	3	5	8	10	13	15												
9	1	4	6	9	12	15	18	21											
10	1	4	7	11	14	17	20	24	27										
11	1	5	8	12	16	19	23	27	31	34									
12	2	5	9	13	17	21	26	30	34	38	42								
13	2	6	10	15	19	24	28	33	37	42	47	51							
14	3	7	11	16	21	26	31	36	41	46	51	56	61						
15	3	7	12	18	23	28	33	39	44	50	55	61	66	72					
16	3	8	14	19	25	30	36	42	48	54	60	65	71	77	83				
17	3	9	15	20	26	33	39	45	51	57	64	70	77	83	89	96			
18	4	9	16	22	28	35	41	48	55	61	68	75	82	88	95	102	109		
19	4	10	17	23	30	37	44	51	58	65	72	80	87	94	101	109	116	123	
20	4	11	18	25	32	39	47	54	62	69	77	84	92	100	107	115	123	130	138

$p=0.01$

$n_1 \backslash n_2$	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
5	-	-	0	1															
6	-	-	1	2	3														
7	-	0	1	3	4	6													
8	-	0	2	4	6	7	9												
9	-	1	3	5	7	9	11	14											
10	-	1	3	6	8	11	13	16	19										
11	-	1	4	7	9	12	15	18	22	25									
12	-	2	5	8	11	14	17	21	24	28	31								
13	0	2	5	9	12	16	20	23	27	31	35	39							
14	0	2	6	10	13	17	22	26	30	34	38	43	47						
15	0	3	7	11	15	19	24	28	33	37	42	47	51	56					
16	0	3	7	12	16	21	26	31	36	41	46	51	56	61	66				
17	0	4	8	13	18	23	28	33	38	44	49	55	60	66	71	77			
18	0	4	9	14	19	24	30	36	41	47	53	59	65	70	76	82	88		
19	1	4	9	15	20	26	32	38	44	50	56	63	69	75	82	88	94	101	
20	1	5	10	16	22	28	34	40	47	53	60	67	73	80	87	93	100	107	114



**Таблица 4**

**Критические значения Т-критерия Вилкоксона  
для уровней статистической значимости  $p=0,05$  и  $p=0,01$   
и объёма выборки  $5 \leq n \leq 50$**

n	p=0.05	p=0.01	n	p=0.05	p=0.01
5	0	-	28	130	101
6	2	-	29	140	110
7	3	0	30	151	120
8	5	1	31	163	130
9	8	3	32	175	140
10	10	5	33	187	151
11	13	7	34	200	162
12	17	9	35	213	173
13	21	12	36	227	185
14	25	15	37	241	198
15	30	19	38	256	211
16	35	23	39	271	224
17	41	27	40	286	238
18	47	32	41	302	252
19	53	37	42	319	266
20	60	43	43	336	281
21	67	49	44	353	296
22	75	55	45	371	312
23	83	62	46	389	328
24	92	69	47	407	345
25	100	76	48	426	362
26	110	84	49	446	379
27	119	92	50	466	397

**Таблица 5**

**Критические значения L-критерия тенденций Пейджа для уровней статистической значимости  $p=0,05$ ,  $p=0,01$  и  $p=0,001$  количества условий  $3 \leq c \leq 6$ , объёма выборки  $2 \leq n \leq 12$**

n	c				
	3	4	5	6	p
2	-	-	109	178	0,001
	-	60	106	173	0,01
	28	58	103	166	0,05
3	-	89	160	250	0,001
	42	87	155	252	0,01
	41	84	150	244	0,05
4	56	117	210	341	0,001
	55	114	204	331	0,01
	54	111	197	321	0,05
5	70	145	259	420	0,001
	68	141	251	409	0,01
	66	137	244	397	0,05
6	83	172	307	499	0,001
	81	167	299	486	0,01
	79	163	291	474	0,05
7	96	198	355	577	0,001
	93	193	346	563	0,01
	91	189	338	550	0,05
8	109	225	403	655	0,001
	106	220	393	640	0,01
	104	214	384	625	0,05
9	121	252	451	733	0,001
	119	246	441	717	0,01
	116	240	431	701	0,05
10	134	278	499	811	0,001
	131	272	487	793	0,01
	128	266	477	777	0,05
11	147	305	546	888	0,001
	144	298	534	869	0,01
	141	292	523	852	0,05
12	160	331	593	965	0,001
	156	324	581	946	0,01
	153	317	570	928	0,05

**Таблица 6**  
**Критические значения критерия  $\chi^2$**   
**для уровней статистической значимости  $p=0,05$  и  $p=0,01$**   
**и количестве степеней свободы  $v=(k-1)(c-1)$**

<b>v</b>	<b>p=0.05</b>	<b>p=0.01</b>	<b>v</b>	<b>p=0.05</b>	<b>p=0.01</b>	<b>v</b>	<b>p=0.05</b>	<b>p=0.01</b>
1	3,84	6,64	35	49,80	57,34	69	89,39	99,23
2	5,99	9,21	36	51,00	58,62	70	90,63	100,43
3	7,82	11,35	37	52,19	59,89	71	91,67	101,62
4	9,49	13,28	38	53,38	61,16	72	92,1	102,82
5	11,07	15,09	39	54,57	62,43	73	93,95	104,01
6	12,59	16,81	40	55,76	63,69	74	95,08	105,20
7	14,07	18,48	41	56,94	64,95	75	96,22	106,39
8	15,51	20,09	42	58,12	66,21	76	97,35	107,58
9	16,92	21,67	43	59,30	67,46	77	98,48	108,77
10	18,31	23,21	44	60,48	68,71	78	99,62	109,96
11	19,68	24,73	45	61,66	69,96	79	100,75	111,14
12	21,03	26,22	46	62,83	71,20	80	101,88	112,33
13	22,36	27,69	47	64,00	72,44	81	103,01	113,51
14	23,69	29,14	48	65,17	73,68	82	104,14	114,70
15	25,00	30,58	49	66,34	74,92	83	105,27	115,88
16	26,30	32,00	50	67,51	76,15	84	106,40	117,06
17	27,59	33,41	51	68,67	77,39	85	107,52	118,24
18	28,87	34,81	52	69,83	78,62	86	108,65	119,41
19	30,14	36,19	53	70,99	79,84	87	109,77	120,59
20	31,41	37,57	54	72,15	81,07	88	110,90	121,77
21	32,67	38,93	55	73,31	82,29	89	112,02	122,94
22	33,92	40,29	56	74,47	83,51	90	113,15	124,12
23	35,17	41,64	57	75,62	84,73	91	114,27	125,29
24	36,42	42,98	58	76,78	85,95	92	115,39	126,46
25	37,65	44,31	59	77,93	87,17	93	116,51	127,63
26	38,89	45,64	60	79,08	88,38	94	117,6	128,80
27	40,11	46,96	61	80,23	89,59	95	118,75	129,97
28	41,34	48,28	62	81,38	90,80	96	119,87	131,14
29	42,56	49,59	63	82,53	92,01	97	120,99	132,31
30	43,77	50,89	64	83,68	93,22	98	122,11	133,48
31	44,99	52,19	65	84,82	94,42	99	123,23	134,64
32	46,19	53,49	66	85,97	95,63	100	124,34	135,81
33	47,40	54,78	67	87,11	96,83			
34	48,60	56,06	68	88,25	98,03			

**Таблица 7**  
**Критические значения критерия Колмогорова-Смирнова**  
**при сопоставлении эмпирического распределения с теоретическим**

n	$d_{\max}$		n	$d_{\max}$	
	p=0,05	p = 0,01		p = 0,05	p = 0,01
5	0,6074	0,7279	50	0,1921	0,2302
10	0,4295	0,5147	60	0,1753	0,2101
15	0,3507	0,4202	70	0,1623	0,1945
20	0,3037	0,3639	80	0,1518	0,1820
25	0,2716	0,3255	90	0,1432	
30	0,2480	0,2972	100	0,1358	
40	0,2147	0,2574	n>100	$1,36/\sqrt{n}$	$1,63/\sqrt{n}$

**Таблица 8**  
**Уровни статистической значимости значений критерия  $\varphi^*$  –**  
**углового преобразования Фишера**

$p \leq$	Последний десятичный знак									
	0	1	2	3	4	5	6	7	8	9
0,00	2,91	2,81	2,70	2,62	2,55	2,49	2,44	2,39	2,35	
0,01	2,31	2,28	2,25	2,22	2,19	2,16	2,14	2,11	2,09	2,07
0,02	2,05	2,03	2,01	1,99	1,97	1,96	1,94	1,92	1,91	1,89
0,03	1,88	1,86	1,85	1,84	1,82	1,81	1,80	1,79	1,77	1,76
0,04	1,75	1,74	1,73	1,72	1,71	1,70	1,68	1,67	1,66	1,65
0,05	1,64	1,64	1,63	1,62	1,61	1,60	1,59	1,58	1,57	1,56
0,06	1,56	1,55	1,54	1,53	1,52	1,52	1,51	1,50	1,49	1,48
0,07	1,48	1,47	1,46	1,46	1,45	1,44	1,43	1,43	1,42	1,41
0,08	1,41	1,40	1,39	1,39	1,38	1,37	1,37	1,36	1,36	1,35
0,09	1,34	1,34	1,33	1,32	1,32	1,31	1,31	1,30	1,30	1,29
0,10	1,29									

**Таблица 9**

**Критические значения коэффициента корреляции  $r_{xy}$  Пирсона для  
уровней статистической значимости  $p=0,05$  и  $p=0,01$   
и количества степеней свободы  $k=n-2$**

k=n-2	p=0.05	p=0.01	k=n-2	p=0.05	p=0.01	k=n-2	p=0.05	p=0.01
5	0.75	0.87	19	0.43	0.55	45	0.29	0.37
6	0.71	0.83	20	0.42	0.54	50	0.27	0.35
7	0.67	0.80	21	0.41	0.53	60	0.25	0.33
8	0.63	0.77	22	0.40	0.52	70	0.23	0.30
9	0.60	0.74	23	0.40	0.51	80	0.22	0.28
10	0.58	0.71	24	0.39	0.50	90	0.21	0.27
11	0.55	0.68	25	0.38	0.49	100	0.20	0.25
12	0.53	0.66	26	0.37	0.48	125	0.17	0.23
13	0.51	0.64	27	0.37	0.47	150	0.16	0.21
14	0.50	0.62	28	0.36	0.46	200	0.14	0.18
15	0.48	0.61	29	0.36	0.46	300	0.11	0.15
16	0.47	0.59	30	0.35	0.45	500	0.09	0.12
17	0.46	0.58	35	0.33	0.42	700	0.07	0.10
18	0.44	0.56	40	0.30	0.39	1000	0.06	0.09

**Таблица 10**

**Критические значения коэффициента ранговой корреляции  $r_s$   
Спирмена для уровней статистической значимости  $p=0,05$  и  $p=0,01$**

n	p=0.05	p=0.01	n	p=0.05	p=0.01	n	p=0.05	p=0.01
5	0,94	—	17	0,48	0,62	29	0,37	0,48
6	0,85	—	18	0,47	0,60	30	0,36	0,47
7	0,78	0,94	19	0,46	0,58	31	0,36	0,46
8	0,72	0,88	20	0,45	0,57	32	0,36	0,45
9	0,68	0,83	21	0,44	0,56	33	0,34	0,45
10	0,64	0,79	22	0,43	0,54	34	0,34	0,44
11	0,61	0,76	23	0,42	0,53	35	0,33	0,43
12	0,58	0,73	24	0,41	0,52	36	0,33	0,43
13	0,56	0,70	25	0,40	0,51	37	0,33	0,43
14	0,54	0,68	26	0,39	0,50	38	0,32	0,41
15	0,52	0,66	27	0,38	0,49	39	0,32	0,41
16	0,50	0,64	28	0,38	0,48	40	0,31	0,40